
Ústav informatiky Akademie věd ČR
Institute of Computer Science, Academy of Sciences of the Czech Republic
Proceedings of the XI. PhD. Conference
Edited by F. Hakl
Doktorandský den ’06
Monínec – Sedlec-Prčice
September 20–22, 2006

Doktorandský den ’06

Ústav informatiky
Akademie věd České republiky

Monínec, Sedlec–Prčice

20.– 22. září 2006

matfyzpress
vydavatelství Matematicko-fyzikální fakulty
Univerzity Karlovy v Praze

Ústav Informatiky AV ČR, Pod Vodárenskou věží 2, 182 07 Praha 8

ISBN 80-86732-87-8

Contents

Doktorandský den Ústavu informatiky Akademie věd České republiky se koná již pojedenácté, nepřetržitě od roku 1996. Tento seminář poskytuje doktorandům, podílejícím se na odborných aktivitách Ústavu informatiky, možnost prezentovat výsledky jejich odborného studia. Současně poskytuje prostor pro oponentní připomínky k přednášené tematice a použité metodologii práce ze strany přítomné odborné komunity.

Z jiného úhlu pohledu, toto setkání doktorandů podává průřezovou informaci o odborném rozsahu pedagogických aktivit, které jsou realizovány na pracovištích či za spoluúčasti Ústavu informatiky AV ČR.

Jednotlivé příspěvky sborníku jsou uspořádány podle jmen autorů. Uspořádání podle tematického zaměření nepovažujeme za účelné, vzhledem k rozmanitosti jednotlivých témat.

Vedení Ústavu informatiky jakožto organizátor doktorandského dne věří, že toto setkání mladých doktorandů, jejich školitelů a ostatní odborné veřejnosti povede ke zkvalitnění celého procesu doktorandského studia zajišťovaného v součinnosti s Ústavem informatiky a v neposlední řadě k navázání a vyhledání nových odborných kontaktů.

1. září 2006

Relative interpretations over first-order fuzzy logic

Post-Graduate Student:
MGR. LIBOR BĚHOUNEK
Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic
behounek@cs.cas.cz

Supervisor:
DOC. PHDR. PETR JIRKŮ, CSC.
Faculty of Philosophy
Charles University in Prague
Celetná 20
116 42 Praha 1
Czech Republic
petr.jirku@ff.cuni.cz

Field of Study:
Logic

This work was supported by grant No. 401/03/H047 of the Grant Agency of the Czech Republic *Logical foundations of semantics and knowledge representation.* The co-advisor for my research in the area of fuzzy logic is Prof. RNDr. Petr Hájek, DrSc.

Abstract

The classical notion of relative interpretation (also known as direct syntactic model) is adapted for multi-sorted first-order fuzzy logics. The level of generality is chosen to suit the needs of its applications in Fuzzy Class Theory.

In formal logic, relative interpretations are a powerful tool that can be used not only for the proofs of relative consistency, but also for direct syntactic constructions of notions of one theory in another. Here we adapt the notion for fuzzy logic and show the analogs of key classical metatheorems. These results allow using relative interpretations of fuzzy theories in essentially the same way they are used in classical metamathematics.

Relative interpretations can be defined at varying levels of generality, the price for greater generality being more preconditions in theorems on invariance under an interpretation. The level of generality chosen here follows the needs of the paper [3]. For relative interpretations see [9]; we follow and slightly generalize the exposition given in [8].

Multi-sorted first-order fuzzy logic with subsumption of sorts has been introduced in [1] for the logic ŁΠ [6, 4]. It is nevertheless obvious that the definitions and proofs of [1] work over any fuzzy logic that axiomatically expands MTL or MTLΔ [5]. In what follows, by "fuzzy logic" we shall therefore mean any logic that in this sense contains MTL; we shall only require that all of its propositional connectives be extensional w.r.t. provable equivalence (otherwise some of the metatheorems below could fail). Crisp identity is assumed in the first-order fuzzy logic under consideration; in models it is always realized as the identity of elements and it can be axiomatized e.g. by the axioms of reflexivity $x = x$ and intersubstitutivity *salva veritate* $x = y \to (\varphi(x) \leftrightarrow \varphi(y))$ for any formula φ (for details see [1]).

Besides the theorems of first-order MTL that are listed in [5], we shall need a few more (meta)lemmata. The following lemma shows that it is possible to bind only some occurrences of a term in the existentialization of a formula:

Lemma 1 *Let $\varphi(x, y)$ be a formula and t a term substitutable for both x and y in φ. Then $\varphi(t, t) \to (\exists x)\varphi(x, t)$.*

Proof: Directly by existentialization on x in φ. QED

Lemma 2 *For an arbitrary term t substitutable for x in $\varphi(x)$ it is provable that*

$$\varphi(t) \quad \leftrightarrow \quad (\forall x)(x = t \rightarrow \varphi(x)) \tag{1}$$
$$\varphi(t) \quad \leftrightarrow \quad (\exists x)(x = t \,\&\, \varphi(x)) \tag{2}$$

Proof: (1) Left to right: from the identity axiom $\varphi(t) \rightarrow (x = t \rightarrow \varphi(x))$ by generalization on x and shifting the quantifier. Right to left: by specification of x to t.

(2) Left to right: $\varphi(t)$ implies $t = t \,\&\, \varphi(t)$, which by Lemma 1 implies $(\exists x)(x = t \,\&\, \varphi(x))$. Right to left: from the identity axiom $x = t \,\&\, \varphi(x) \rightarrow \varphi(t)$ by generalization on x and shifting the quantifier to the antecedent. QED

Corollary 3 *Any formula is equivalent to a formula in which logical functions are applied only to variables and occur only in atomic subformulae of the form $y = F(x_1, \ldots, x_k)$.*

Proof: Using Lemma 2, inductively decompose nested terms $s(t)$ by $\varphi(s(t)) \leftrightarrow (\exists x)(x = t \,\&\, \varphi(s(x)))$ and finally by $\varphi(F(x_1, \ldots, x_k)) \leftrightarrow (\exists y)(y = F(x_1, \ldots, x_k) \,\&\, \varphi(y))$ for all F. QED

Recall from [1] that the language of a multi-sorted first-order fuzzy logic is a quintuple $(\mathbf{S}, \preceq, \mathbf{P}, \mathbf{F}, \mathbf{A})$, where $\mathbf{S}$ is a non-empty set of sorts, $\preceq$ is a partial ordering of $\mathbf{S}$ indicating the subsumption of sorts, $\mathbf{P}$ and $\mathbf{F}$ are disjoint sets of predicate resp. function symbols, and $\mathbf{A}$ is an arity function that assigns a finite sequence of sorts to each element of $\mathbf{P} \cup \mathbf{F}$ (the sequence must be non-empty for elements of $\mathbf{F}$). If $P \in \mathbf{P}$ and $\mathbf{A}(P) = (s_1, \ldots, s_k)$, then $P(t_1, \ldots, t_k)$ is a well-formed atomic formula iff the term t_i is of sort s_i for all $i = 1, \ldots, k$. If $\mathbf{A}(F) = (s_1, \ldots, s_k, s_{k+1})$, then $F(t_1, \ldots, t_k)$ is a well-formed term, of sort s_{k+1}, iff the term t_i is of sort s_i for all $i = 1, \ldots, k$. For more details on multi-sorted first-order fuzzy logics see [1, §2.2].

Definition 4 (Interpretation of a language) *Let $\mathbf{L} = (\mathbf{S}, \preceq, \mathbf{P}, \mathbf{F}, \mathbf{A})$ and $\mathbf{L}' = (\mathbf{S}', \preceq', \mathbf{P}', \mathbf{F}', \mathbf{A}')$ be two multi-sorted first-order languages. An* interpretation of the language $\mathbf{L}$ in the language $\mathbf{L}'$ *is a (metamathematical) mapping $\star$ which assigns to each sort $s \in \mathbf{S}$ a function symbol $F^\star_s \in \mathbf{F}'$ of arity $\mathbf{A}'(F^\star_s) = (s_\star, s^\star)$ for some $s_\star, s^\star \in \mathbf{S}'$, to each predicate symbol $P \in \mathbf{P}$ a predicate symbol $P^\star \in \mathbf{P}'$, and to each function symbol $F \in \mathbf{F}$ a function symbol $F^\star \in \mathbf{F}'$, and which satisfies the following conditions:*

- *For all $s, r \in \mathbf{S}$, if $s \preceq r$ then $s^\star \preceq r^\star$.*
- *For all $P \in \mathbf{P}$, if $\mathbf{A}(P) = (s_1, \ldots, s_k)$ and $\mathbf{A}'(P^\star) = (r_1, \ldots, r_k)$ then $s^\star_i \preceq r_i$ for all $i = 1, \ldots, k$.*
- *For all $F \in \mathbf{F}$, if $\mathbf{A}(F) = (s_1, \ldots, s_{k+1})$ and $\mathbf{A}'(F^\star) = (r_1, \ldots, r_{k+1})$, then $s^\star_i \preceq r_i$ for all $i = 1, \ldots, k$ and $r_{k+1} \preceq s^\star_{k+1}$.*

An interpretation $\star$ of $\mathbf{L}$ in $\mathbf{L}'$ extends by metamathematical induction on the complexity of terms and formulae of $\mathbf{L}$ to a mapping (also denoted by $\star$) which assigns to each term t of $\mathbf{L}$ a term $t^\star$ of $\mathbf{L}'$ and to each formula φ of $\mathbf{L}$ a formula $\varphi^\star$ of $\mathbf{L}'$ as follows:

- *The i-th variable x^s_i of each sort s in $\mathbf{L}$ is assigned the term $F^\star_s(x^{s_\star}_i)$ of sort $s^\star$, where $x^{s_\star}_i$ is the i-th variable of sort $s_\star$.*
- *Each term $F(t_1, \ldots, t_k)$ of $\mathbf{L}$ is assigned the term $F^\star(t^\star_1, \ldots, t^\star_k)$.*
- *Each formula $P(t_1, \ldots, t_k)$ of $\mathbf{L}$ is assigned the formula $P^\star(t^\star_1, \ldots, t^\star_k)$.*
- *Each formula $x = y$ of $\mathbf{L}$ is assigned the formula $x^\star = y^\star$.*

- *For all k-ary propositional connectives* $\mathbf{c}$*, each formula* $\mathbf{c}(\varphi_1, \ldots, \varphi_k)$ *of* $\mathbf{L}$ *is assigned the formula* $\mathbf{c}(\varphi_1^\star, \ldots, \varphi_k^\star)$.
- *Each formula* $(\forall x^s)\varphi$ *resp.* $(\exists x^s)\varphi$ *of* $\mathbf{L}$ *is assigned the formula* $(\forall x^{s_\star})\varphi^\star$ *resp.* $(\exists x^{s_\star})\varphi^\star$.

Remark 5 Notice that we allow reinterpreting *variables* of sort s by *functions* from $s_\star$ to $s^\star$. This is necessitated by the applications in [3], where we need to interpret variables by functional terms (e.g., when identifying x with the pair $\langle x, 0\rangle$). A straightforward interpretation of a sort s by another sort r is covered by this definition, taking the identity function on sort $s_\star$ for $F_s^\star$ and $s^\star = s_\star = r$.

Remark 6 In Definition 4, the logical symbols (except for variables) are left unaffected by the translation $\star$. The notion of interpretation can be defined more generally to include also the specification of the translations of $[(\forall x)\varphi]^\star$, $[(\exists x)\varphi]^\star$, $(x = y)^\star$, and $[\mathbf{c}(\varphi_1, \ldots, \varphi_k)]^\star$ for each propositional connective $\mathbf{c}$. In Definition 9 we would then require the provability of the interpreted logical axioms and rules.

Notice that in the latter case, the background logic of the interpreted language or theory may be allowed to differ from the background logic of the original language or theory. For empty theories, we then get an interpretation of one logic in another. An example of such kind is the interpretation of the logic $\mathrm{PC}(*)$ of a particular ŁΠ-representable t-norm $*$ in ŁΠ, which takes $\to$ of $\mathrm{PC}(*)$ to $\to_*$ of ŁΠ, $\&$ of $\mathrm{PC}(*)$ to $\&_*$ of ŁΠ, etc. By a recent result (oral presentation by Marchioni and Montagna at IPMU'06), the interpretation is faithful, i.e., $\mathrm{PC}(*) \vdash \varphi$ iff ŁΠ $\vdash \varphi_*$, for any formula φ of $\mathrm{PC}(*)$. Another example of this kind are Gödel-style interpretations, e.g., the $\neg\neg$-interpretation of classical logic in SMTL (or stronger) or the Δ-interpretation of classical logic in MTLΔ (or stronger). (Notice that Gödel-style interpretations require a further generalization of the rule for the interpretation of atomic formulae.) In this paper, however, we shall only use interpretations which leave the logical symbols absolute, and thus do not change the underlying logic.

Definition 7 (Absolute and invariant notions) *Let $\star$ be an interpretation of the language $\mathbf{L}$ in the language $\mathbf{L}'$ and let $\mathbf{T}'$ be a theory in the language $\mathbf{L}'$. Let $\varphi(x_1, \ldots, x_k)$ be a formula of $\mathbf{L}$ and let all non-logical symbols of φ belong to $\mathbf{L}'$ as well. Then the formula φ is called* absolute *(in the theory $\mathbf{T}'$ w.r.t. the interpretation $\star$) iff $\mathbf{T}' \vdash \varphi(x_1^\star, \ldots, x_k^\star) \leftrightarrow \varphi^\star$. Similarly, a predicate P or a functor F is called absolute, if the formula $P(x_1, \ldots, x_k)$ resp. $y = F(x_1, \ldots, x_k)$ is absolute.*

Let furthermore $\mathbf{L}'$ contain the sorts of all variables that occur in φ. Then we will call φ invariant *(in the theory $\mathbf{T}'$ w.r.t. the interpretation $\star$) iff $\mathbf{T}' \vdash \varphi \leftrightarrow \varphi^\star$. A predicate P or a functor F is called invariant, if the formula $P(x_1, \ldots, x_k)$ resp. $y = F(x_1, \ldots, x_k)$ is invariant.*

Observation 8 *If φ is both absolute and invariant w.r.t. $\star$ in $\mathbf{T}'$, then $\mathbf{T}' \vdash \varphi(x_1^\star, \ldots, x_k^\star) \leftrightarrow \varphi(x_1, \ldots, x_k)$.*

Definition 9 (Interpretation of a theory) *Let $\mathbf{T}$ be a theory in the language $\mathbf{L}$ and $\mathbf{T}'$ a theory in the language $\mathbf{L}'$. An interpretation $\star$ of $\mathbf{L}$ in $\mathbf{L}'$ is called an* interpretation of the theory $\mathbf{T}$ in the theory $\mathbf{T}'$ *iff $\mathbf{T}' \vdash \varphi^\star$ for each formula φ which is a logical axiom of identity or an axiom of the theory $\mathbf{T}$.*

The requirement in Definition 9 that the interpreted identity axioms be provable is automatically satisfied if all functions $F_s^\star$ together are injective:

Lemma 10 *Let $\star$ be an interpretation of the language $\mathbf{L}$ in the language $\mathbf{L}'$ and let $\mathbf{T}'$ be a theory in the language $\mathbf{L}'$. If $\mathbf{T}' \vdash F_s^\star(x^{s_\star}) = F_t^\star(y^{t_\star}) \to x^{s_\star} = y^{t_\star}$ for all sorts s, t in $\mathbf{L}$, then the interpreted axioms of identity are provable in $\mathbf{T}'$.*

Proof: The axiom of reflexivity $x = x$ translates into $x^\star = x^\star$, which is an instance of the reflexivity axiom of identity in $\mathbf{T}'$. The intersubstitutivity axiom $x = y \to [\varphi(x) \leftrightarrow \varphi(y)]$ translates into the formula of the form

$$F_s^\star(x^{s_\star}) = F_t^\star(y^{t_\star}) \to [\psi(x^{s_\star}) \leftrightarrow \psi(y^{t_\star})]$$

which is provable in $\mathbf{T}'$ by the assumption of the lemma and the instance for ψ of the intersubstitutivity axiom of $\mathbf{T}'$. QED

The usual theorems on interpretations known from classical logic remain valid for interpretations over fuzzy logics as well. The following theorems give examples of such results.

Observation 11 *A composition of two interpretations (between languages or theories) is an interpretation (between languages or theories, respectively).*

Since the composition is obviously associative and the identical mapping is always an interpretation of a theory in itself, the languages or theories over a fuzzy logic form a category just like in classical logic, allowing categorial constructions on fuzzy theories.

Theorem 12 *Let $\star$ be an interpretation of the theory $\mathbf{T}$ in the theory $\mathbf{T}'$. Then for any formula φ in the language of $\mathbf{T}$, if $\mathbf{T} \vdash \varphi$ then $\mathbf{T}' \vdash \varphi^\star$.*

Proof: By induction on the proof of φ: by the requirement of Definition 9, the interpreted axioms of $\mathbf{T}$ and those of identity are provable in $\mathbf{T}'$, and all other logical axioms and rules are translated by $\star$ again into the instances of logical axioms and rules (observe that the term $t^\star$ is substitutable for $x^{s\star}$ iff t is substitutable for x^s). QED

Definition 13 (Faithful interpretations) *The interpretation $\star$ of the theory $\mathbf{T}$ in the theory $\mathbf{T}'$ is* faithful *iff for all formulae φ in the language of $\mathbf{T}$ it holds that $\mathbf{T} \vdash \varphi$ iff $\mathbf{T}' \vdash \varphi^\star$.*

A faithful interpretation $\star$ of $\mathbf{T}$ in itself such that $\varphi^{\star\star} \equiv \varphi$ is called a duality.

Example 14 (Identical interpretation) *If the theory $\mathbf{T}'$ in the language $\mathbf{L}'$ extends the theory $\mathbf{T}$ in the language $\mathbf{L}$, then the identical interpretation of $\mathbf{L}$ in $\mathbf{L}'$ (i.e., $x^\star = x$, $P^\star = P$, and $F^\star = F$ for all sorts and symbols) interprets $\mathbf{T}$ in $\mathbf{T}'$. The interpretation is faithful iff $\mathbf{T}'$ extends $\mathbf{T}$ conservatively.*

The following lemma gives a method how to prove the faithfulness of an interpretation in some cases.

Lemma 15 *Let $\star$ interpret $\mathbf{T}$ in its extension $\mathbf{T}'$ and let $s_\star = s$ for all sorts in $\mathbf{T}$. Let furthermore*

$$\mathbf{T}' \vdash P^\star(F^\star_{s_1}(x_1^{s_1}), \ldots, F^\star_{s_k}(x_k^{s_k})) \leftrightarrow P(x_1^{s_1}, \ldots, x_k^{s_k}) \tag{3}$$

$$\mathbf{T}' \vdash F^\star_s(y^s) = F^\star(F^\star_{s_1}(x_1^{s_1}), \ldots, F^\star_{s_k}(x_k^{s_k})) \leftrightarrow y^s = F(x_1^{s_1}, \ldots, x_k^{s_k}) \tag{4}$$

for all function symbols F and predicate symbols P in the language of $\mathbf{T}$ (including the identity predicate).

Then $\mathbf{T}' \vdash \varphi^\star \leftrightarrow \varphi$ for all formulae φ in the language of $\mathbf{T}$ (i.e., all notions in the language of $\mathbf{T}$ are invariant under $\star$).

If furthermore $\mathbf{T}'$ extends $\mathbf{T}$ conservatively, then $\star$ is faithful.

Proof: The first claim is proved straightforwardly by induction on the subformulae of φ. By (3), (4) and Corollary 3 we can assume that $\mathbf{T}' \vdash \psi^\star \leftrightarrow \psi$ holds for all atomic subformulae ψ in φ. Propositional combinations preserve the property $\mathbf{T}' \vdash \psi^\star \leftrightarrow \psi$, since our definition of interpretation leaves all propositional connectives absolute and in the logics under consideration all connectives are extensional w.r.t. provable equivalence. For $\psi \equiv (\forall x^s)\chi$, since $s_\star = s$, its translation $\psi^\star$ is $(\forall x^s)\chi^\star$, and thus $\mathbf{T}' \vdash (\forall x^s)\chi^\star \leftrightarrow (\forall x^s)\chi$ follows from the induction hypothesis $\mathbf{T}' \vdash \chi^\star \leftrightarrow \chi$ by the rules of MTL (similarly for $\exists$).

The claim of faithfulness under conservativity: $\mathbf{T}' \vdash \varphi^\star \leftrightarrow \varphi$ entails ($\mathbf{T}' \vdash \varphi^\star$ iff $\mathbf{T}' \vdash \varphi$), and by conservativity $\mathbf{T}' \vdash \varphi$ iff $\mathbf{T} \vdash \varphi$; thus $\mathbf{T}' \vdash \varphi^\star$ iff $\mathbf{T} \vdash \varphi$. QED

Remark 16 Definition 4 requires that all sorts and symbols occurring in the definitions of $x^\star$, $P^\star$, and $F^\star$ be present in the language $\mathbf{L}'$. Following the usual mathematical practice, we shall not distinguish between a theory and its extensions by conservative definitions. Thus we shall allow giving $x^\star$, $P^\star$, and $F^\star$ by the defining formulae or terms for the needed predicates, functors, and sorts, provided the definitions are conservative.

For the conservative introduction of predicate and function symbols see [7]: the definition of a predicate symbol by an axiom $P(x_1,\ldots,x_k) \leftrightarrow \varphi(x_1,\ldots,x_k)$ is conservative and eliminable for any formula φ, while the introduction of a function symbol $F(x_1,\ldots,x_k)$ by an axiom $\varphi(x_1,\ldots,x_k,F(x_1,\ldots,x_k))$ is conservative on condition that $(\exists x_{k+1})\varphi(x_1,\ldots,x_k,x_{k+1})$ is provable in the theory; the definition is eliminable if the uniqueness of such x_{k+1} is provable in the theory. (In multi-sorted languages, the obvious conditions on the sorts of the arguments must be ensured.)

For the definition of sorts, it is easy (but tedious) to check that a sort s subsumed in a sort t can be introduced by an axiom $(\exists x^s)(x^t = x^s) \leftrightarrow \varphi(x^t)$, which is conservative if the theory proves that $(\exists x^t)\varphi(x^t)$ and that φ is *crisp;* if it is further required that $s \preceq s'$ for any sort s', the conservativity is ensured if the theory further proves $\varphi(x^t) \to (\exists x^{s'})(x^t = x^{s'})$.

The apparatus of relative interpretations is widely applicable in all sorts of formal fuzzy theories. Since Fuzzy Class Theory FCT of [1] is proposed in [2] as a foundational theory for fuzzy mathematics, relative interpretations of various fragments of FCT in itself are of special importance. In Example 17 I give an incomplete list of such interpretations (the details will be given in a separate paper). Some of them (e.g., the upper shift or the relativization) prove important (even if often intuitively obvious) metamathematical properties of FCT, while others codify constructions which either obviate some of the syntactic restrictions of FCT (e.g., the singleton shift), or can be useful in various areas of fuzzy mathematics formalized in FCT (e.g., the "$\times\{0\}$" interpretation, employed in [3]).

Example 17 *The following constructions are important interpretations of* FCT *(or some of its fragments) in* FCT*:*

- Identical interpretations. *Propositional fuzzy logic, classical theory of the identity of individuals, the classical theory of identity of tuples, the theory of fuzzy classes, the theory of fuzzy relations, and monadic Henkin-style higher-order fuzzy logic are all fragments of* FCT *given by a suitable restriction of the language (admitting only some sorts of variables). It can be shown that they can be axiomatized by the axioms of* FCT *restricted to the same language with an additional axiom stating that the sorts for tuples do not exhaust the universal sort of the same order.* FCT *extends these theories conservatively, and thus the identical interpretations of the respective fragments represent all of the above theories faithfully in* FCT.
- Upward shift. *The translation* $\sharp$ *that consists in raising the order of all variables by* 1 *is an interpretation of* FCT *in itself (since the axioms of* FCT *are invariant under* $\sharp$*). All definitions and theorems of* FCT *can thus be propagated to all higher orders by iteration of* $\sharp$*.*
- Relativization. *Restricting all quantifiers to a* crisp *class (resp. its iterated crisp powers in higher orders) is an interpretation of* FCT *in itself. The domain of discourse thus can be arbitrarily chosen from some basic universe (as long as it is crisp).*
- Singleton shift. FCT *does not allow classes to contain elements of different orders (e.g.,* $\{x, X\}$*). Nevertheless, they can be simulated by means of faithful interpretations. It can be shown that the interpretation* $\{\cdot\}$ *("singleton shift") which maps* x *to* $\{x\}$ *is a faithful interpretation of the theory of identity (which exhausts the relevant features of atomic elements) in the theory of fuzzy classes. The mixed class* $\{x, X\}$ *thus can be "encoded" by the class* $\{x^{\{\cdot\}}, X\} = \{\{x\}, X\}$*. (Further adjustments can be made in order to make the backward translation one-to-one and make it work at all levels of the type hierarchy.) Thus by this interpretation, mixed classes of arbitrary orders are available in* FCT.

- Transposition. *Switching all pairs $\langle x, y\rangle$ to $\langle y, x\rangle$ is a duality in* FCT. *Dual forms of the theorems on fuzzy relations thus need not be proved (e.g.,* $\mathrm{dom}(A \times B) = A$*, follows from* $\mathrm{rng}(A \times B) = B$*).*
- Relational representation of classes. *Fuzzy classes can be represented among fuzzy relations by identifying atomic elements x with pairs $\langle x, 0\rangle$ (for a fixed element 0); any fuzzy class A is then identified with the fuzzy relation $A \times \{0\}$. This interpretation is employed in [3] for proving hosts of theorems on fuzzy relations and classes at once.*

Remark 18 The interpretations of Example 17 often state an "isomorphism" of some structures in FCT. The need of using interpretations arises primarily from the fact that the notion of isomorphism (not even a bijection) has not yet been developed inside FCT. (Since all notions in FCT are in general fuzzy, this notion would need a careful analysis.) Nevertheless, since FCT is a formal syntactic theory, the metamathematical apparatus of interpretations is very suitable for such tasks, and the "syntactic isomorphisms" obtained by the method of interpretation are usually easier to prove than they would be inside the theory.

References

[1] L. Běhounek and P. Cintula, "Fuzzy class theory," *Fuzzy Sets and Systems*, vol. 154, no. 1, pp. 34–55, 2005.

[2] L. Běhounek and P. Cintula, "From fuzzy logic to fuzzy mathematics: A methodological manifesto," *Fuzzy Sets and Systems*, vol. 157, no. 5, pp. 642–646, 2006.

[3] L. Běhounek and M. Daňková, "Relational compositions in Fuzzy Class Theory." Under preparation, 2006.

[4] P. Cintula, "The ŁΠ and ŁΠ$\frac{1}{2}$ propositional and predicate logics," *Fuzzy Sets and Systems*, vol. 124, no. 3, pp. 289–302, 2001.

[5] F. Esteva and L. Godo, "Monoidal t-norm based logic: Towards a logic for left-continuous t-norms," *Fuzzy Sets and Systems*, vol. 124, no. 3, pp. 271–288, 2001.

[6] F. Esteva, L. Godo, and F. Montagna, "The ŁΠ and ŁΠ$\frac{1}{2}$ logics: Two complete fuzzy systems joining Łukasiewicz and product logics," *Archive for Mathematical Logic*, vol. 40, no. 1, pp. 39–67, 2001.

[7] P. Hájek, "Function symbols in fuzzy logic," in *Proceedings of the East-West Fuzzy Colloquium*, (Zittau/Görlitz), pp. 2–8, IPM, 2000.

[8] A. Sochor, *Klasická matematická logika*. Prague: Karolinum Press, 2001. In Czech.

[9] A. Tarski, A. Mostowski, and R. M. Robinson, *Undecidable Theories*. Amsterdam: North-Holland, 1953.

Vizualizace dat v relační databázi s využitím teorie formálních jazyků

doktorand:
ING. ZUZANA CAPEKOVÁ
ÚI AVČR, Pod Vodárenskou věží 2, 182 07 Praha 8
FM TUL, Hálková 6, 461 17 Liberec 17
capekova@cs.cas.cz, zuzana.capekova@tul.cz

školitel:
ING. JÚLIUS ŠTULLER, CSC.
ÚI AVČR, Pod Vodárenskou věží 2, 182 07 Praha 8
stuller@cs.cas.cz

obor studia:
Technická kybernetika

Abstrakt

Článek se zabývá možností efektivního návrhu webových stránek pro vizualizaci dat v relačních databázích s využitím teorie formálních jazyků a překladů. Jazyk závorkových výrazů libovolné hloubky vnoření patří mezi bezkontextové jazyky. Jestliže omezíme hloubku vnoření závorek zvolenou konstantou, lze pro takový jazyk sestrojit regulární gramatiku, která jej generuje. Generování věty regulárního jazyka pomocí pravidel pravé regulární gramatiky se vyznačuje ”koncovou rekurzí” tj. větná forma obsahuje nejvýše jediný neterminální symbol, který je zároveň symbolem větné formy nejvíce vpravo. Výběr pravidla pro expanzi neterminálu je možné řídit informacemi získanými z relační databáze. Koncovou rekurzi lze snadno vyjádřit pomocí iterace. V dostatečně dlouhých větách všech bezkontextových jazyků lze na určitých místech zopakovat podřetězec dané věty tak, aby získaná věta patřila do jazyka. Kontextové jazyky podobnou vlastnost obecně nemají, nicméně i tam můžeme v izolovaných případech zopakovat na určitých místech části věty s identickým výsledkem. Pro získávání dat z relačních databází je možné definovat množinu SQL dotazů uspořádanou do stromu pomocí ”master-detail” referenčních vztahů. Na příkladu jazyka HTML ve spojení se závorkovou strukturou reprezentující strom dotazů konečné výšky je demonstrována možnost dynamického generování www stránek výhradně na základě deklarativní konfigurace s vyloučením programátorské práce.

1. Úvod

Současný rychlý rozvoj internetu klade zvýšené nároky na tvůrce webových stránek. Dnes již není možné spoléhat se pouze na statický webový obsah. Návrh www stránek v běžně užívaných systémech dynamického generování stránek (php, asp, jsp, Java-servlet apod.) však může úspěšně provádět jen odborník alespoň zhruba znalý problematiky imperativního a objektového programování. V této práci představujeme jeden ze způsobů, jak umožnit visualizaci dat z relační databáze pomocí dynamického generování www stránek výhradně na základě deklarativní konfigurace s možností dodatečně upravovat grafickou úpravu generovaných stránek s vyloučením programátorské práce. Dynamické generování HTML stránek můžeme považovat za speciální případ generování řetězců složených se symbolů alfanumerické abecedy. Zaměření této práce proto není omezeno čistě jen na generování HTML stránek, naopak s výjimkou konkrétních příkladů v HTML jsou všechny závěry práce aplikovatelné i na generování xml nebo rtf dokumentů. Prezentaci dat v podobě rtf dokumentu v kombinaci s široce rozšířeným editorem rtf dokumentů (Microsoft Word) lze též s výhodou využít jako alternativní systém tisku uživatelsky modifikovatelných tiskových sestav. Základním principem postupu prezentovaného v této práci je obohacení HTML (resp. rtf atd.) dokumentu o závorkovou strukturu reprezentující algoritmus získávání dat z relační databáze. S výhodou též využíváme vlastností jazyka HTML resp. rtf, zejména se jedná o možnost 0..n krát zopakovat ve vybraných větách jazyka určité podřetězce dané věty tak, že výsledná posloupnost symbolů zůstává větou jazyka. Dynamické generování dokumentů na základě uvedeného principu je možné výrazně zefektivnit, pakliže provedeme předzpracování dokumentu. Předzpracování dokumentu je možné provést s využitím teorie formálních překladů. Výsledkem předzpracování je binární vnitřní forma, která je vhodnější pro opakované generování dokumentu. Práce využívá výsledky teorie relačních databází [1, 5] a teorie formálních jazyků a

překladů [1, 2, 3]. Zápis gramatiky resp. grafické vyjádření konečných automatů je uvedeno v notaci podle [1, 2]. Přehled základních pojmů je uveden v [1, 2, 3].

2. Strom dotazů

Pro účely této práce definujeme strom dotazů nad množinou relací takto:

Definition 1 *N*

echť **R** je množina relací. Strom dotazů budiž orientovaný stromový graf ve kterém platí:

1. Hrany jsou orientovány ve směru od rodičovského uzlu k potomkům.
2. Každý uzel stromu je ohodnocen jedinou relací $r \in R$.
3. Každá hrana směřující od předka u_i ohodnoceného relací $r_i = \langle A_i, D_i, T_i \rangle$ k potomkovi u_j ohodnocenému relací $r_j = \langle A_j, D_j, T_j \rangle$je ohodnocena funkcí $f_j : T_i \times T_j \longrightarrow \{0, 1\}$. V souladu s rozšířenou definicí operace selekce v relační algebře podle[5], je funkce *f* tvořena logickým výrazem sestrojeným z logických spojek a porovnání hodnot atributů.

Stromový graf se vyznačuje m.j. tím, že počet jeho hran je vždy o jedna menší než počet uzlů a že od rodičovského uzlu k některému z přímých potomků vede vždy jen jedna jediná hrana. Můžeme proto označit uzly, hrany a jim přiřazené relace resp. funkce pomocí celých čísel takto:

- Uzly jsou označeny jednoznačně celými čísly *1..n*, kde *n* je počet uzlů grafu.
- Kořen stromu je označen číslem *1*.
- Hrana *h*, která vede od rodiče u_i k potomkovi u_j je označena číslem *j*.
- Symbolem r_i označíme relaci, kterou je ohodnocen uzel u_i.
- Symbolem f_j označíme funkci, kterou je ohodnocena hrana stromu h_j.
- Funkce f_j, kterou je ohodnocena hrana stromu h_j směřující od předka u_i k potomkovi u_j, určuje pro každou kombinaci prvků z relace r_i a r_j jejich obsahovou souvislost.
- Vstupem funkce je vždy jeden prvek z relace r_i a jeden prvek z relace r_j.
- Výstupem funkce je číslo *0*, pakliže vstupující prvky obsahově nesouvisí. V opačném případě funkce vrátí číslo *1*.

2.1. Traverzování stromu dotazu

Strom dotazů umožňuje zorganizovat pořadí zpracování prvků relací, kterými jsou ohodnoceny uzly stromu. Předpokládejme, že v průběhu traverzování je strom dotazů neměnný, tj. má konstantní množinu uzlů, množinu hran i incidenční zobrazení. Traverzování stromu a zpracování prvků relací nechť probíhá takto:

Pro každý prvek *t* relace r_i, omezené selekcí na podmínku $f_i = 1$, se navštíví všichni přímí následníci uzlu u_i ve stanoveném neměnném pořadí. Návštěva přímého potomka u_j je spojena s přechodem po hraně h_j ve směru její orientace a s provedením selekce na relaci r_j podle podmínky definované takto: $f_j(t, x) = 1$.

Prvek $t \in T_i$ představuje aktuálně zpracovávaný prvek z relace předka a je konstantní v průběhu provádění selekce, symbol *x* označuje po řadě jednotlivé prvky relace r_j, tj. platí $x \in T_j$. Pakliže je relace získaná selekcí prázdná, provede se návrat do rodičovského uzlu. V opačném případě se postup rekurzivně zopakuje.

V případě kořene u_1 se postupuje obdobně s tím rozdílem, že neproběhne selekce na relaci r_1 a po ukončení probírky všech prvků relace r_1 algoritmus skončí.

Uvedený postup umožňuje zpracovat (např.vypsat) všechny prvky relací přiřazených následníkům ("detail") obsahově související s právě zpracovávaným prvkem relace přiřazené rodičovskému uzlu ("master"). Až poté dojde ke zpracování dalšího prvku relace přiřazené rodičovskému uzlu. V okamžiku provedení selekce je jisté, že relace na úrovni rodičovského uzlu není prázdná, tj. selekční kritérium má zajištěny korektní vstupní hodnoty.

2.2. Vyjádření stromu dotazů pomocí jazyka závorkových výrazů

Pro vyjádření stromu dotazů pomocí jazyka závorkových výrazů můžeme využít atributované terminální symboly otevírací a uzavírací závorky a syntaxi jazyka formálně vyjádřit pomocí atributové gramatiky. Atributová gramatika je čtveřice *AG=(G,A,V,F)*, kde *G=(N,T,P,S)* je základní gramatika atributové gramatiky, *A* je množina atributů, *V* je zobrazení přiřazující každému symbolu z $N \bigcup T$ množinu jeho atributů, F je množina sémantických pravidel [2].

Předpokládejme, že strom dotazů obsahuje vždy alespoň jeden uzel. Pak můžeme sestrojit atributovou gramatiku takto:

1. Základní gramatika G = (N, T, P, S)
   $T = \{[,]\}$
   $N = \{S, Z\}$
   $P = \{S \longrightarrow [Z, Z \longrightarrow SZ, Z \longrightarrow]\}$
2. Množina atributů
   $A = \{node\_number\}$
3. Zobrazení V
   $V = \{([, \{node\_number\}), (], \{node\_number\}), (S, \{\}), (Z, \{\})\}$
4. Množina F - prázdná
   $F = \{\}$

Převod stromu dotazů na větu uvedeného jazyka provedeme rekurzivním průchodem přes uzly stromu. Při první návštěvě uzlu zapisujeme otevírací závorku s atributem nastaveným na číslo uzlu, pak rekurzivně projdeme potomky a před návratem k rodiči zapíšeme uzavírající závorku s atributem nastaveným na číslo uzlu. Takto získaný výraz označme pro účely tohoto textu symbolem "**VSD**". Uvedené vyjádření stromu dotazů není jednoznačné v tom smyslu, že pro kompletní zpětný převod je třeba ještě doplnit převodní zobrazení, které na základě čísla uzlu určí uzlu příslušející relaci a hraně (přicházející od rodiče) přiřadí příslušející funkci. Dále je pro tvar **VSD** výrazu podstatné pořadí procházení potomků. Neméně důležitý je převod zadaného řetězce na strom dotazů resp. kontrola zadaného atributovaného řetězce a zjištění, zda odpovídá zadanému stromu dotazů. Tento úkol řeší lexikální a syntaktická analýza. Snadno ověříme, že základní gramatika námi sestrojené atributové gramatiky je bezkontextová a navíc LL(1). Proto můžeme snadno sestrojit deterministický syntaktický analyzátor pomocí techniky rekurzivního sestupu [1, 2].

3. Jazyk závorkových výrazů

Zabývejme se nyní izolovaně jazykem závorkových výrazů *L*, který je obecně bezkontextový. Vyznačuje se existencí dvou symbolů v abecedě – levé a pravé závorky. V každé větě jazyka je celkově stejný počet pravých a levých závorek a navíc v libovolném prefixu věty je počet levých závorek větší nebo roven počtu pravých závorek. Jazyk závorek je např. generován gramatikou $G = (N, T, P, S)$

- $T = \{[,]\}$

- $N = \{S, Z\}$
- $P = \{S \longrightarrow \varepsilon, S \longrightarrow ZS, Z \longrightarrow [S]\}$

Uvedená gramatika *G* generuje nekonečný jazyk $L(G) = L$.

Symbolem $h(s)$ označme maximální hloubku vnoření závorek ve větě $s \in L(G)$, přičemž hloubku vnoření závorky definujeme obvyklým rekurzivním předpisem $h = h\_parent + 1$ hloubka závorky nikde nevnořené je 1, $h(\varepsilon) = 0$. Hloubka vnoření závorek je v každé izolované větě omezena, neboť věta jazyka má vždy konečný počet symbolů. Pro libovolné celé číslo $n \in N$ však vždy existuje věta $s \in L(G)$, taková že $h(s) = n$.

Zabývejme se nyní jazykem $\{L2 = s : s \in L \wedge h(s) \leq k\ \}$, kde $k \in N$ e zvolená konstanta. Jazyk *L2* je tvořen podmnožinou řetězců jazyka *L*, kde je maximální hloubka vnoření závorek $h(s)$ všech řetězců $s \in L2$ omezena zvolenou konstantou $k \in N$. Jazyk L2 patří mezi regulární jazyky a můžeme pro něj sestrojit konečný automat, který jej přijímá, resp. regulární gramatiku, která jej generuje.

I v případě, že povolíme více typů závorek, tj. uvažujeme více než jednu levou závorku a ke každé levé závorce přiřadíme jí odpovídající pravou závorku, dostaneme při konstantou omezené hloubce vnoření regulární jazyk popsatelný např. konečným automatem.

Vzpomeňme nyní na atributovaný závorkový výraz **VSD** z kapitoly 2.2, který jsme použili jako jednu z možností pro vyjádření stromu dotazů. Jelikož má konečnou hloubku, má smysl zabývat se sestrojením konečného automatu resp. regulární gramatiky generující jazyk, jehož věty jsou odvozeny od **VSD** výrazu konečným (i nulovým) počtem opakování závorky na daném místě v těsném sousedství s ostatními výskyty. Atribut "node_number" považujme nyní za typové označení závorek. Hledaný jazyk je tvořen řetězci nad abecedou $A = \{$"$[1''$, "$1]''$, "$[2''$, "$2]''\ldots$ "$[n''$, "$n]''\}$, kde hodnoty atributu se staly součástí symbolu abecedy. Dále požadujeme, aby závorku příslušející uzlu u_j (vyskytuje se ve výrazu jen jednou) bylo možno libovolně-krát zopakovat na daném místě v rámci závorky příslušející rodiči uzlu u_j.

Konečný automat přijímající uvedený jazyk sestrojíme na základě znalosti struktury stromu dotazů a požadovaného řazení potomků uzlů stromu, obojí můžeme jednoznačně určit z **VSD**. Požadovaný deterministický konečný automat MSD=(Q, T, δ, q_0, F) sestrojíme takto:

1. Množina vnitřních stavů Q:
   $Q = \{q\_0, q\_left\_1, q\_right\_1, \ldots.., q\_left\_n, q\_right\_n\}$
   Každému uzlu stromu dotazů příslušejí dva stavy – "left" a "righ". Navíc je zde stav q_0, který je počátečním stavem automatu.

2. Vstupní abeceda T:
   $T = A = \{$"$[1''$, "$1]''$, "$[2''$, "$2]''\ldots$ "$[n''$, "$n]''\}$

3. Přechodové zobrazení δ:
   $\delta(q\_0,$ "$[1'') = q\_left\_1$
   $\delta(q\_left\_i,$ "$i]'') = q\_right\_i$ pro $i \in 1..n$
   $\delta(q\_right\_i,$ "$[i'') = q\_left\_i$ pro $i \in 1..n$
   $\delta(q\_left\_i,$ "$[j'') = q\_left\_j$, je- li uzel u_j přímým potomkem uzlu u_i
   $\delta(q\_right\_j,$ "$i]'') = q\_right\_i$ je- li uzel u_j přímým potomkem uzlu u_i
   Nakonec vyřešíme přechody mezi potomky stejného rodiče: Potomky jednoho rodiče seřadíme v pořadí podle výskytu ve **VSD** výrazu a z pravého stavu předchůdce vedeme přechod do levého stavu všech následníků, přechod spojíme se čtením symbolu levé závorky s číslem příslušného následníka "[descendant_number".

4. Počáteční stav je q_0

5. Množina koncových stavů F:
   $F = \{q\_0, q\_prava\_1\}$
   Koncové stavy představují uzavření závorky příslušející kořeni stromu dotazů resp. přijetí prázdného řetězec znamenající nulový výskyt závorky přiřazené kořeni stromu dotazů.

Výraz **VSD** označující strom dotazů můžeme tedy převést na konečný automat a posléze na regulární gramatiku generující jazyk požadovaných vlastností.
S výhodou nyní užijeme skutečnosti, že atributovaný závorkový výraz **VSD** nepopisuje strom dotazů jednoznačně. Může existovat více stromů dotazů se stejným vyjádřením pomocí **VSD** výrazu - stromy se shodným **VSD** výrazem se mohou lišit v přiřazení relací uzlům nebo v definici funkcí přiřazených hranám. Zkusme nyní sestrojit algoritmus pro generování vět jazyka **L(MSD)**, vznik věty jazyka bude při tom řízen informacemi získanými z traversování stromu dotazů. Na základě znalosti **MSD** automatu sestrojíme nejprve pravou regulární gramatiku postupem uvedeným v [1]. Gramatika takto sestavená velice názorně kopíruje strukturu zdrojového konečného automatu.

Generování věty v pravé regulární gramatice probíhá jako posloupnost přímých derivací z počátečního symbolu gramatiky. Přímá derivace v pravé regulární gramatice znamená nahrazení jediného neterminálu X větné formy pravou stranou pravidla, jehož levou stranu tvoří právě neterminál X. Pravidel se stejným neterminálem na levé straně může být v pravé regulární gramatice několik, proto je třeba umět se při generování věty rozhodnout pro jedno určité pravidlo.

Při expanzi neterminálů vytvořeným z "levých" stavů (q_left_i), jsou na výběr pravé strany obsahující po řadě neterminály vytvořené z "levých" stavů potomků uzlu u_i a "pravého" stavu uzlu u_i. Analogicky při expanzi neterminálů vytvořeným z"pravých" stavů (q_right_i), jsou na výběr pravé strany obsahující po řadě neterminály vytvořené z "levích" stavů následníků uzlu u_i v "sourozenecké řadě", "levého" stavu uzlu u_i a "pravého" stavu rodičovského uzlu.

Výběrovým kritériem budiž existence resp. neexistence dosud nezpracovaných prvků v relaci příslušející některému ze zainteresovaných uzlů. Konkrétní postup budiž tento:

Při expanzi neterminálu q_left_i e již znám právě zpracovávaný prvek *t* relace r_i. Proto je možné určit selekci na realce potomků uzlu u_i.

Příslušné pravé strany pravidel představují přechod ke generování obsahu na základě potomků uzlu u_i resp. přechod na další prvek relace r_i. Pravidlo obsahující na pravé straně neterminál q_left_j se užije v případě, že výsledek selekce na relaci u_i není prázdná relace. V opačném případě se provede přechod na další pravidlo v pořadí určeném **VSD** výrazem a testování probíhá stejně. Při expanzi neterminálu q_right_i e nejprve označí prvek *t* za zpracovaný a testuje existence dalšího dosud nezpracovaného prvku v relaci r_i. Pakliže takový prvek existuje, použije se pravidlo q_right_i $\longrightarrow$ "[i"q_left_i, jinak se ve známém pořadí zkoumají pravidla vedoucí k levým závorkám sourozenců, pravidlo vedoucí k pravé závorce předka se užije v případě, že ostatní možnosti nelze užít.

Snadno zjistíme, že při řízení expanze neterminálů výše navrženým způsobem kopíruje generování věty jazyka **L(MSD)** traverzování stromu dotazů. Operační složitost je tedy dána nároky na traverzování plus nároky na generování věty. **VSD** výraz při tom m.j. určuje pořadí procházení potomků resp. pořadí kontroly použitelnosti pravidel pro expanzi příslušného neterminálu. Pakliže takto navrženému algoritmu necháme zpracovat jiný strom dotazů pro daný **VSD** výraz, vygeneruje algoritmus obecně jinou větu jazyka **L(MSD)**. Pořadí provádění kontrol a expanzí je neměnné pro všechny povolené vstupní stromy dotazů. Proto je můžeme s výhodou zakódovat např. do posloupnosti operací, která se provede v okamžiku generování výstupu.

Na závěr prozkoumejme ještě možnost umístit mezi závorky vstupního **VSD** výrazu nějaké pro vlastní generování "balastní" symboly. Závorku pak zopakujeme na daném místě včetně jejího obsahu. Úprava **MSD** automatu je relativně jednoduchá – přibudou lineární úseky stavů a přechodů generující dané konstantní řetězce výplňových symbolů, **MSD** automat zůstane deterministický. Algoritmus generování vět jazyka

L(MSD) zůstane též v podstatě nezměněn, žádné dodatečné rozhodování není potřeba.

3.1. Grafické formátování informací

Nyní se zaměříme blíže na řetězce terminálních symbolů umístěných mezi závorkami VSD výrazu. V dostatečně dlouhých větách všech bezkontextových jazyků lze na určitých místech zopakovat podřetězec dané věty tak, aby získaná věta patřila do jazyka [3]. Kontextové jazyky podobnou vlastnost obecně nemají, nicméně i tam můžeme v izolovaných případech zopakovat na určitých místech části věty s identickým výsledkem. Na problematiku můžeme nahlédnout i z druhé strany, tj. existuje možnost vložit atributované závorky **VSD** výrazu do věty jiného jazyka tak, aby repetice obsahu závorky na daném místě a následné vypuštění **VSD** závorek neporušilo příslušnost věty k jazyku. Toho můžeme s výhodou využít např. ve spojení s jazykem HTML takto:

1. HTML dokument obohatíme o atributované závorky.
2. Pomocí stromu dotazu modifikujeme generování vstupní věty.
3. Z výstupné věty odstraníme atributované závorky, čímž se z ní stane opět věta jazyka HTML.

Jediné úskalí postupu je v bodu 1. – rozmístění atributovaných závorek po původní větě jazyka nemůže být zcela libovolné. Rozmístování závorek po původní větě jazyka však může být přenecháno k tomu určenému editoru. Jako příklad již existující implementace uveďme jazyk RTF a editor Microsoft Word – Microsoft Word obsahuje funkci “Vložit záložku”, která vloží atributovanou pravou a levou závorku požadovaných vlastností. Úplně postačuje funkce umožňující korektní vložení izolovaných závorek, vzájemný vztah závorek se vyřeší v rámci předzpracování dokumentu.

4. Realizace generátoru dokumentů

Doposud jsme probrali jen základní “motor” generátoru dokumentů. Od reálného exportního systému požadujeme ještě výstup hodnot atributů prvků relací do výsledného textu. V průběhu traverzování stromu dotazů jsou při průchodu uzlu u_i známy a konstantní aktuálně zpracovávané prvky relací všech předchůdců uzlu u_i až ke kořeni stromu. Je proto přirozené umožnit zápis hodnot atributů prvku *t* relace r_i v rámci celé závorky příslušející uzlu u_i ve **VSD** výrazu. Definujme nyní pojem “**SCH (SD)**”.

Definition 2 *S*

chéma **SCH(SD)** stromu dotazů **SD** je stromový graf se stejnou množinou uzlů *a* hran *a* se shodným incidenčním zobrazením jako zdrojový strom dotazů **SD**. Uzly schématu jsou ohodnoceny schémata relací, hrany schématu nejsou ohodnoceny.

Na základě znalosti schématu stromu dotazů je možné určit všechny možné tvary **VSD** výrazu a ověřit rozmístění atributů mezi závorkami **VSD** výrazu. Kompletní atributová gramatika definující syntaxi rozeznávanou syntaktickým analyzátorem vznikne jednoduchým rozšířením gramatiky pro **VSD** výraz o atributovaný symbol *a*“atributu relace” a atributovaný symbol *p*“vmezeřený text”. Na jejím základě lze sestrojit syntaktický analyzátor a prekladač vstupního textu do vhodné binární vnitřní formy jak je uvedeno např. [2]. Schéma stromu dotazů musí být zadáno spolu se vstupním textem. Překladač jej využije při ověření umístění symbolů značících atributy relací v rámci příslušných závorek atd. Interpret vnitřní formy (vlastní generátor dokumentů) je možné realizovat s využitím poznatků z kapitoly 2 a 3 tohoto textu, hodnoty atributů je třeba v okamžiku zápisu konvertovat do textové podoby.

4.1. Příklad užití ve spojení s jazykem HTML

Předpokládejme, že máme k dispozici implementovaný překladač a implementovaný interpret vnitřní formy. Lexikální analyzátor v překladači nechť předpokládá zápis atributovaných závorek pomocí lexikálních elementů ve tvaru `<%BS_nodeid%>` pro otevírací závorku a `<%BE_nodeid%>` pro uzavírací závorku. Zápis označení atributů se předpokládá ve tvaru `<%variableid%>`. Předpokládejme nyní, že chceme zobrazit izolovanou tabulku v určitém grafickém formátování, data získáme z relační databáze v podobě jediné relace – strom dotazů má pouze kořenový uzel.

Example 1 *V*

zorový dokument obohacený o závorky a atributy bude vypadat např. takto:

```
<html><body> <table>
    <tr><td>First name</td><td>Last name</td></tr>
    <%BS_1%>
    <tr><td><%txtfirstname%></td><td><%txtlastname%></td></tr>
    <%BE_1%>
</table> </body></html>
```

Případná změna resp. doplnění formátování může probíhat bez zásahu programátora s např. pomocí vhodného HTML editoru.

Example 2 *V*

ýsledný dokument bez závorek a s nahrazenými atributy bude vypadat např. takto:

```
<html> <body> <table>
    <tr><td>First name</td><td>Last name</td></tr>
    <tr><td>Zuzana</td><td>Capekova</td></tr>
<tr><td>Iva</td><td>Zlamalova</td></tr> </table> </body></html>
```

5. Závěr

Použití uvedeného postupu není omezeno jen na HTML jazyk. Postup lze aplikovat i ve spojení s jazykem xml, rtf apod. Editace vzhledu dokumentů může být přenechána zcela neškoleným uživatelům. Např. v případě jazyka rtf je možné pro změnu vzhledu dokumentu přímo použít aplikaci MS Word (funkce ”vložit záložku” vytvoří závorky požadovaných vlastností), v důsledku čehož není nutné uživatele přeškolovat na specifická vývojová prostředí, návrháře reportů apod. Předzpracování vzorového dokumentu do vhodné binární vnitřní formy navíc zvyšuje efektivitu metody a minimalizuje výpočetní i paměťové nároky na generování dokumentu.

Literatura

[1] Melichar, B. “Jazyky a překlady”, Vydavatelství ČVUT 1999.

[2] Müller, K. “Programovací jazyky”, Vydavatelství ČVUT 2002.

[3] Mareš, J. “Jazyky, gramatiky a automaty”, Vydavatelství ČVUT 2004,

[4] Stuller, J. “Přednášky z databázových systémů”, `http://www.cs.cas/ stuller`.

[5] Maier, D. “The theory of relational databases”, Pitman 1983.

Frailty Models in Survival Analysis

Post-Graduate Student:
MGR. VÁCLAV FALTUS, M.SC.
Department of Medical Informatics
Instutite of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic
faltus@euromise.cz

Supervisor:
DOC. ZDENĚK VALENTA, M.SC., PH.D.
Department of Medical Informatics
Instutite of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic
valenta@euromise.cz

Field of Study:
Biomedical Informatics

The work was suported by the grant 1M06014 of the Ministry of Education of the Czech Republic

Abstract

The aim of this paper is to present an overview of the methods used in survival analysis and especially modeling survival data. Since the topic of my future Ph.D. thesis is *Statistical models for correlated survival data* we introduce the use of frailties as an equivalent of random effects in common statistical modeling together with its connection to correlation. Frailty model, how model with frailties is called, uses frailties as a parameter for individuals. Those who are most frail will experience an event earlier than others.

Keywords: survival analysis, frailty models

1. Introduction

Survival analysis is analysis of time to the occurrence of an event. Examples of such data arise in many diverse fields, such as medicine, biology, public health, epidemiology, engineering, economics, and demography. In epidemiology the event is not always a death, but it can be, for example, a first occurrence of relapse, duration of response to treatment, time to development of a disease, duration of stay in hospital, and duration of a seizure. Survival analysis attempts to answer questions such as: what is the fraction of a population which will survive past a certain time? Of those that survive, at what rate will they die or fail? Can multiple causes of death or failure be taken into account? How do particular circumstances or characteristics increase or decrease the probability of survival?

Analysis of survival data is very often complicated by the issue of censoring, where the event is known only to not have occurred during a certain period of time and by truncation, where the individuals enter the study only if they survive a sufficient length of time or individuals are included in the study only if the event has occurred by a given time. These two issues play a very important role in estimating model parameters and in modeling survival data. The theoretical basis for the analysis of survival data has been solidified by connecting it to the study of counting processes and martingale theory [6]. This theory, among many other benefits, substantially helped with accounting for censoring.

In this paper, we start with an introduction to the key concepts in survival analysis: the hazard, survival, and cumulative hazard functions. Then we turn to regressions models with emphasis on frailty model as an extension of the Cox proportional hazards model. The frailty model may be also called the shared frailty models because the frailty may be parameter shared among members of one group or family. We also discuss some models used in multivariate survival data and show the universality of the concept of frailties.

2. Basic concepts

Random variables analyzed in survival analysis are called the failure times T and they are always non-negative $T \geq 0$. T can either be discrete (taking a finite set of values, e.g. $a_1, a_2, \ldots, a_n$) or continuous

(defined on $[0, \infty)$).

The concept used in describing time-to-event phenomena is the survival function, the probability of an individual surviving beyond time t. It is defined as

$$S(t) = \mathrm{P}(T > t).$$

Survival function is non-increasing function with a value of 1 at the origin and 0 at infinity. When T is a continuous random variable then $S(t)$ is a non-increasing continuous function. When T is a continuous variable, the survival function is a complement of the cumulative distribution function, $S(t) = 1 - F(t)$, where $F(t) = 1 - P(T \leq t)$. The survival function is the integral of the probability density function $f(x)$. Then

$$S(t) = 1 - F(t) = 1 - P(T \leq t) = \int_t^\infty f(x)dx, \tag{1}$$

and

$$f(x) = -\frac{dS(x)}{dx}.$$

Fundamental in survival analysis is the hazard function. It is also known as the conditional failure rate in reliability, the force of mortality in demography, the intensity function in stochastic processes, the age-specific failure rate in epidemiology or simply as the hazard rate. The hazard rate is defined by

$$\begin{aligned} h(t) &= \lim_{\Delta t \to 0} \frac{\mathrm{P}(T \in [t, t+\Delta t) | T \geq t)}{\Delta t} = \\ &= \lim_{\Delta t \to 0} \frac{\mathrm{P}(T \in [t, t+\Delta t))}{\Delta t \cdot \mathrm{P}(T \geq t)} = \\ &= \lim_{\Delta t \to 0} \frac{F(t+\Delta t) - F(t)}{\Delta t \cdot S(t)} = \frac{F'(t)}{S(t)}. \end{aligned} \tag{2}$$

Then

$$h(t) = \frac{F'(t)}{S(t)} = \frac{(1 - S(t))'}{S(t)} = \frac{-S'(t)}{S(t)} = -\ln'(S(t)). \tag{3}$$

A related quantity is the cumulative hazard function, defined by

$$H(t) = \int_{y=0}^{t} h(y)dy = -\ln S(t) + \ln S(0) = -\ln(S(t)). \tag{4}$$

Then

$$S(t) = e^{-H(t)}.$$

3. Modeling survival data

A problem frequently encountered in analyzing survival data is that of adjusting the survival function to account for concomitant information (sometimes referred to as covariates, explanatory variables or independent variables).

A popular concept for modeling the relationship of covariates to a survival outcomes is represented by the Cox proportional hazards model. Let X_{ij} be the covariate of the ith person, where $i = 1, \ldots, n$ and $j = 1, \ldots, p$. The set of covariates then, like in linear regression, forms an $n \times p$ matrix and X_i is used to denote the covariate vector for subject i (the ith row of the matrix).

The Cox model specifies the hazard for individual i as

$$h_i(t) = h_0(t) \cdot e^{X_i\beta}, \tag{5}$$

where h_0 is unspecified nonnegative function of time called the baseline hazard, and β is $p \times 1$ column vector of coefficients. Event rates cannot be negative (observed deaths cannot unhappen), and the exponential thus plays the an important role in ensuring that the final estimates are a physical possibility.

Because the hazard ratio for two subjects with fixed covariate vectors X_i and X_j,

$$\frac{h_i(t)}{h_j(t)} = \frac{h_0(t) \cdot e^{X_i\beta}}{h_0(t) \cdot e^{X_j\beta}} = \frac{e^{X_i\beta}}{e^{X_j\beta}} = e^{[X_i - X_j]\beta} \tag{6}$$

is constant over time, the model is also known as the proportional hazards model. Estimations of β is based on the partial likelihood function introduced by Cox. Commonly used algorithm for finding approximations to the roots of real-valued function is the Newton-Raphson algorithm. The idea of the method is that one starts with a value which is reasonably close to the true root, then replaces the function by its tangent and computes the root of this tangent. This root of the tangent will typically be a better approximation to the function's root. The procedure is iterative and compares the nth guess with $n + 1$th, with the initial value usually set to 0. This algorithm is robust for the Cox partial likelihood. Convergence problem are very rare and easily addressed by simple methods such as step-halving.

3.1. The concept of frailty

In the last several years there has been active research in the area of survival models involving random effects. In this setting, a random effect is a continuous variable, that describes excess of risk or frailty for distinct categories, such as individuals or families. The idea is that individuals are characterized by frailty parameter that reflects their state of health. Those who are most frail will die earlier than the others.

Frailty models are also used in making adjustments for overdispersion in survival studies. Here, the frailty represents the total effect on survival of the covariates not measured when collecting information on individual subjects. If these effects are ignored, the resulting survival estimates may be misleading. Corrections for overdispersion allow for adjustments for other important effects that were unaccounted for in the study.

Computationally, frailty parameters are usually viewed as unobserved covariates. This leads to the use of EM (expectation-maximization) algorithm as an estimation tool. EM algorithm is an algorithm for obtaining maximum likelihood estimates of parameters in probabilistic models, where the model depends on unobserved latent variables. EM alternates between performing an expectation (E) step, which computes an expectation of the likelihood by including the latent variables as if they were observed, and a maximization (M) step, which computes the maximum likelihood estimates of the parameters by maximizing the expected likelihood obtained at the E step. The parameters obtained at the M step are then used to begin another E step, and the process is repeated. However, the algorithm is slow, proper variance estimates require further computation, and no implementation appears in any of the more widely available packages. The computation can be approached instead as a penalized Cox model.

The most common model used is a shared frailty model. As we will see, it is an extension of the proportional hazards. Assume that each subject $i, i = 1, \ldots, n$, is a member of a single group $j, j = 1, \ldots, q$. Then we write the proportional hazards model

$$h_i(t) = h_0(t) \cdot e^{(X_i\beta + Z_i\omega)}, \tag{7}$$

where X_i and Z_i are the ith rows of covariate matrices $X_{n\times p}$ and $Z_{n\times p}$ respectively. X and β correspond to p fixed effects in the model, ω is a vector containing the q unknown random effects or frailties, and Z is a design matrix – Z_{ij} equals 1 if subject i is a member of family j, 0 otherwise.

For subject i, which is a member of the jth family the proportional hazards model can be also written as

$$h_{i(j)}(t) = h_0(t) \cdot w_j \cdot e^{X_i\beta}, \tag{8}$$

where the w_j is the frailty parameter for family j. This can be easily rewritten in the form of equation 7, with i ranging over all subjects,

$$w_j = e^{\omega_j},$$

and Z a matrix of indicator variables such that $Z_{ij} = 1$ if subject i is a member of family j and 0 otherwise. In this model, each individual can belong to only one family. The frailty parameter w_j has multiplicative effect on the hazard rates.

According to a penalty function and the choice of a design matrix we distinguish between gamma frailty model and Gaussian random effects model. The advantage of frailty having either of these two distributions is that the shared frailty model can be written exactly as a penalized likelihood and further estimation can be done more easily. Additional assumptions are discussed below.

Penalty function

$$p(\omega) = (1/\omega) \cdot \sum (\omega_i - e^{\omega_i})$$

and the design matrix Z as defined for shared frailty model gives us the equivalence with gamma frailty model. The ω_is are distributed as the logs of iid (independent identically distributed) random variables and the tuning parameter θ is their variance. For this frailty distribution, the correlation of subject within groups is $\theta/(2+\theta)$.

Penalty function

$$p(\omega) = (1/2\omega) \cdot \sum \omega_i^2$$

and general design matrix Z gives the Gaussian random effects model. The tuning parameter θ of the penalty function is the variance of the ω_is.

The fact that both penalties are the log-likelihoods for a random sample of $\omega_1, \ldots, \omega_n$ from the appropriate distribution raises the question whether other frailty distributions can be accommodated within penalized framework. From the viewpoint of the penalized fitting procedure θ is as nuisance or "tuning" parameter of the computation.

The variance can be a fixed parameter set by the user, either set directly or chosen indirectly by specifying the degrees of freedom for the random effects term. Next option is to seek an overall best variance by minimizing the AIC (Akaike's Information Criterion) or corrected AIC. The idea behind the AIC is to examine the complexity of the model together with goodness of its fit to the sample data, and to produce a measure which balances between the two. The formula for computing AIC involves the likelihood and number of parameters of examined model. The corrected AIC uses much larger penalty as the number of parameters approaches the sample size n. This leads to models which are a bit more conservative in the amount of model parameters.

For the gamma frailty, the profile likelihood for θ can be used. This gives a global solution that is identical [6] to the EM algorithm for a shared gamma frailty model. When a Gaussian distribution of the frailty is chosen, the variance θ of the random effect can be chosen based on an approximate REML (restricted maximum likelihood) equation. Restricted maximum likelihood serves to estimate the variance of any distribution and is also based on the likelihood principle.

3.2. Shared Gamma frailty models

As it was said, the most commonly used estimation procedure in frailty models is the EM algorithm. It gives discrete estimator of the distribution and does not allow direct estimation of the hazard function. Rondeaou et al. [2] present how to use maximum penalized likelihood estimation in estimating continuous hazard function in a shared gamma-frailty model with right-censored and left-truncated data. Instead of penalizing the frailties [6] the hazard function is penalized.

Let us consider the model in which the hazard function will partly depend on an unobservable random variable thought to act multiplicatively on the hazard and allow for stratum-specific baseline hazards. For

the jth individual ($j = 1, \ldots, n_{ih}$) of the hth stratum ($h = 1, \ldots, K$) and the ith group ($i = 1, \ldots, G$), let T_{ihj} denote the survival times under study. We also assume the independence of survival times and left-truncation times. The censoring times are assumed independent of the failure times and of the frailties Z_i.

Then the hazard function conditional on the frailty is

$$h_{ihj}(t|Z_i) = Z_i \cdot h_{0h}(t) \cdot e^{X_{ihj}\beta}, \tag{9}$$

where $h_{0h}(t)$ is the baseline hazard function for stratum h; $X_{ihj} = (X_{1ihj}, X_{2ihj}, \ldots, X_{pihj})$ denotes the covariate row vector for the j^{th} individual of stratum h and group i, and β is the corresponding vector of regression parameters.

Conditionally on the frailty Z_i, the failure times $T_{ih1}, T_{ih2}, \ldots, T_{ihn_{ih}}$ are assumed to be independent. It is also assumed that the Z_i's are independently and identically distributed from a gamma distribution with mean 1 and unknown variance θ. The probability density function is then

$$g(z) = \frac{z^{(1/\theta-1)} \cdot e^{-z/\theta}}{\Gamma(1/\theta) \cdot \theta^{1/\theta}}. \tag{10}$$

Large values of θ signify a closer positive relationship between the subjects of the same group and greater heterogeneity among the groups.

Instead of the use of the Cox partial likelihood the full likelihood obtained by integrating out the frailty Z_i from the joint likelihood is presented. Because of allowance for left-truncated data the likelihood is expressed conditional on involved failure times. Rondeau et al. [2] give the exact formula of the full log-likelihood which can be directly maximized to obtain the estimates of β, θ and the baseline hazard function $h_0(t)$, which is assumed to be smooth. Maximum penalized estimators (MPnLE) of $h_0(t)$, β and θ are then obtained by maximizing function consisting of the full log-likelihood and the penalty. The penalty has large values for rough functions and contains also smoothing parameter specific for particular stratum. The idea is that the penalty is for large values of the smoothing parameter forced toward zero and if the smoothing parameter is small, then the main contribution to penalized likelihood will be the full log-likelihood mentioned above.

Rondeau et al. also show the way how to obtain estimators of the baseline hazard function. It is an approximation obtained by using splines. Splines are polynomial functions which are combined linearly to give the interpolation of a function. In addition, confidence bands for the baseline hazard estimators can be given using the variance estimates. The choice of smoothing parameter can be done heuristically by plotting several curves and by choosing that which seems most realistic.

3.3. Nested Gamma frailty models

The frailty model is a random effect survival model, which allows for unobserved heterogeneity or for statistical dependence between observed survival data. The nested frailty model accounts for the hierarchical clustering of the data by including two nested random effects. They substitute the frailty Z_i from equation 9, only the structure is a bit more complicated by the fact that one frailty is nested into another one which refers to higher hierarchical level. Thus nested frailty models are particularly appropriate when data are clustered at several hierarchical levels and it does not matter if the situation arised naturally or by design of the study. In such cases it is important to estimate the parameters of interest as accurately as possible by taking into account the hierarchical structure of the data. Rondeau et. al [3]. present a maximum penalized likelihood estimation (MPnLE) to estimate non-parametrically a continuous hazard function in a nested gamma-frailty model with right-censored and left-truncated data. The estimators for the regression coefficients and the variance components of the random effects are obtained simultaneously.

To demonstrate the MPnLE method and the nested frailty model two examples were given [3]. One was for modeling the effect of particulate air pollution on mortality in different areas with two levels of geographical regrouping. The other application concerned recurrent infection times of patients from different hospitals.

Simulation study was performed and it proved that the semi-parametric approach yields satisfactory results. Using a shared frailty model instead of nested frailty model with two levels of regrouping leads to inaccurate estimates, with an overestimation of the variance of the random effects. Next, even when the frailty effects are fairly small in magnitude, they are important since they alter the results in a systematic pattern.

Let us consider a multilevel proportional hazards model with two sets of nested random effects that act multiplicatively on the hazard, so that a large value of these variables increases the hazard. There are several types of models for expressing how the hazard function depends on the explanatory variables. The most popular model when analyzing epidemiological survival data is the proportional hazards model even if some other models as additive or accelerated failure time can be used. Proportional hazards models are semi-parametric and fairly flexible and their covariates can be time-dependent. We treat the case of a cohort with G independent clusters $(i = 1, \ldots, G)$. Within the ith cluster, there are J_i correlated sub-clusters $(j = 1, \ldots, J_i)$. We treat the case of right-censored and left-truncated data. T_{ijk} denotes the survival times under study for subject k $(k = 1, \ldots, K_{ij})$ from subgroup j, and group i. We also assume the independence of survival times T_{ijk} and left-truncation times.

Then we define two random effects v_i and w_{ij} and assume that the cluster-level random effects v_i and the sub-cluster random effects w_{ij} are independent and gamma-distributed random effects $(\Gamma(1/\alpha; 1/\alpha))$ and $(\Gamma(1/\eta; 1/\eta))$ with $E(v_i) = 1, var(v_i) = \alpha$ and $E(w_{ij}) = 1, var(w_{ij}) = \eta$.

For identifiability, they have a mean equal to 1 at birth (i.e. at duration $t_{ijk} = 0$). If the variance is null, then observations from the same group are independent. A larger variance implies greater heterogeneity in frailty across groups and a greater correlation of the survival times for individuals belonging to the same group. Mainly for reasons of mathematical convenience, the frailty terms are often assumed to follow a gamma distribution. Gamma distribution is chosen because of its correspondence to shared frailty model and the possibility to express the shared frailty model as penalized likelihood. To construct the likelihood function, apart from the usual assumption of independent censoring, the censoring must be non-informative for v_i and w_{ij}.

The hazard function conditional on the two frailties v_i and w_{ij}, for individual (i, j, k) is

$$h_{ijk}(t|v_i, w_{ij}) = v_i \cdot w_{ij} \cdot h_0(t) \cdot e^{X_{ijk}\beta}, \tag{11}$$

where $h_0(t)$ is the baseline hazard function; $X_{ijk} = (X_{1ijk}, \ldots, X_{pijk})$ denotes the covariate row vector for the kth individual, with p the number of covariates, and β is the corresponding vector of regression parameters.

Rondeau et al. estimated the baseline hazard function in a nested frailty model non-parametrically using a penalized full likelihood assuming both left-truncated and right-censored survival data. This makes it possible to estimate simultaneously the regression coefficients, the variance component parameters and especially a smooth hazard function, which cannot be correctly estimated using conventional non-parametric methods.

A major advantage of nested frailty models is their ability to analyze data that are also correlated at several different hierarchical levels. The eventual hierarchical structure of the data needs to be taken into account in survival analysis to obtain accurate inferences. Ignoring random cluster effects may result in overlooking the importance of certain cluster effects and calls into question the validity of traditional statistical techniques such as the shared frailty model. The nested frailty model proved helpful in diagnosing the source of correlation in data. This method can be easily extended to bivariate frailty models, making it possible to treat two events simultaneously per subject.

3.4. Multivariate survival data

The shared frailty model can also be viewed as a specific kind of the common risk model. The frailty is the term that describes the common risk, acting as a factor on the hazard function. In most settings the common risks are assumed random and therefore we can speak about mixture model. Given the frailties the observations are assumed independent. This is then called conditional independence.

The bivariate survival function can be defined as $S(t_1, t_2) = P(T_1 > t_1, T_2 > t_2)$ and in the multivariate case obviously $S(t_1, t_2, \ldots, t_n) = P(T_1 > t_1, T_2 > t_2, \ldots, T_n > t_n)$. P. Hougaard [7] shows also the forms of the bivariate survival function by assuming certain distributions of the frailty. It is also shown that the dependence can be assessed by by the rank-based correlation-type measures like Kendall's τ or Spearman's ρ and these measures depend only on the frailty distribution. That is, they are independent of the hazard function.

In many cases, we may need extensions of the common risk models discussed above and, similarly, more models with varying degrees of dependence. The general frailty approach can be used to create a random treatment by group interaction or other models with several sources of variation. Nested model discussed by Rondeau et al. may be seen as special univariate case of multivariate models discussed in [7]. It was already mentioned above that the two levels in nested model can be extended. This situation can arise in the field of genetics where there are several sources of variation present. For twins the question may be whether the dependence is the same for monozygotic and dizygotic twins. P. Hougaard shows also an example of the nested trivariate parallel data model, which is a model that would be applicable for a sibling group where individuals 1 and 2 are twins and individual 3 is a single birth. Survival times of the first two siblings T_1, T_2 may be strongly dependent. T_3 may show more modest dependence. This scenario can be accomplished by three frailties. In general, each frailty may have different distribution. P. Hougaard proposes the use of semi-parametric estimate. That requires extension for a program for the shared frailty models and seems to be complicated.

So far we discussed multiplicative frailty models but additive models are plausible too, mainly because the multiplicative models are not able to handle all dependence structures. Therefore, the additive models seem to be more operational. The disadvantage of the additive model is that more parameters are needed to define the models where the parameters cannot be identified from subsets of the data.

Frailty can be also modeled as stochastic process. P. Hougaard discusses independent increments frailty models, piecewise gamma model, moving average model, hidden cause of death model and the Woodbury-Manton model. All these models apply when there is a instantaneous and/or short-term frailty. The ideas are given, however the mathematical complexity of these models is a major problem. Instantaneous dependence has a mathematical appeal allowing for many theoretical evaluations. It is obtained by using a frailty varying randomly over time, where the hazard is described by the independent increments of the process. Short-term dependence models seem to be a more relevant extension of the frailty model and from mathematical point of view, there seems to be a lot of potential. The key problem is the computational burden. Models with with time-varying frailties seem particularly important for recurrent events data, but most such model treat only additive piecewise constant frailties and with a limited number of intervals. [7]

4. Conclusion

Short overview of frailty models used in survival analysis was given together with discussion and references to available literature. Discussion concerning censoring, truncation, estimation of parameters using the likelihood, partial likelihood and other methods together with numerical procedures exceeds the scope of this paper and thus were only slightly mentioned. The use and rising popularity of frailty models used in analysis of correlated censored and possibly truncated data is relatively new and active research in this field is taking place. In section about multivariate survival data there is a bit of a discussion about correlation and conditional independence. In my prospective Ph.D. thesis I would like to focus on statistical aspects of models involving frailty parameters and this paper was in my opinion a good start to accomplish that goal.

References

[1] C. A. McGilchrist and C. W. Aisbett, "Regression with Frailty in Survival Analysis", *Biometrics*, vol. 47, pp. 461–466, 1991.

[2] Virginie Rondeau, Daniel Commenges, and Pierre Joly, "Maximum Penalized Likelihood Estimation in a Gamma-Frailty Model" *Lifetime Data Analysis*, vol. 9, pp. 139–153, 2003.

[3] Virginie Rondeau, Daniel Commenges, and Pierre Joly, "Nested frailty models using maximum penalized likelihood estimation" *Statistics in Medicine*, 2006.

[4] Erik Parner, "Asymptotic Theory for the Correlated Gamma-Frailty Model", *The Annals of Statistics*, vol. 26(1), pp. 183–214, 1998.

[5] John P. Klein and Melvin L. Moeschberger, "Survival Analysis: Techniques for Censored and Truncated Data", *Springer*, 1997.

[6] Terry M. Therneau and Patricia M. Grambsch, "Modelling Survival Data: Extending the Cox Model", *Springer*, 2000.

[7] Philip Hougaard "Analysis of Multivariate Survival Data", *Springer*, 2000.

[8] Vladimír Bencko, Karel Hrach, Marek Malý, Hynek Pikhart, Jindra Reissigová, Štěpán Svačina, Marie Tomečková, and Jana Zvárová, "Statistické metody v epidemiologii: svazek 2", *Nakladatelství Karolinum*, 2003.

On a maximum attainable accuracy of some segregated techniques for saddle point problems

Post-Graduate Student:
ING. PAVEL JIRÁNEK
Department of Modelling of Processes
Faculty of Mechatronics, TU Liberec
Hálkova 6
461 17 Liberec
Czech Republic
pavel.jiranek@tul.cz

Supervisor:
DOC. DR. ING. MIROSLAV ROZLOŽNÍK
Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic
miro@cs.cas.cz

Field of Study:
Science engineering

This work was supported by the MSMT CR under the project 1M0554 “Advanced Remedial Technologies”.

Abstract

In this contribution we analyze numerical behavior of some iterative methods used for solving large-scale saddle point problems. Two representatives of such methods are the Schur complement reduction method and the null-space projection method. They are both based on the transformation of the original problem to the reduced form which is solved iteratively (e.g. by the steepest descent or the conjugate gradient methods etc.) giving one block component of the solution vector. The remaining unknowns are solved by back substitution from the original system. Depending on the actual implementation, we estimate the maximum attainable accuracy of the computed approximate solution.

1. Introduction

We consider the saddle point system with the 2×2 block structure

$$\mathcal{A} \begin{pmatrix} x \\ y \end{pmatrix} \equiv \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \tag{1}$$

where the block $A \in \mathbb{R}^{n \times n}$ is symmetric positive definite, the block $B \in \mathbb{R}^{n \times m}$ ($m \leq n$) has rank m. There is a large selection of solution techniques used to solve the systems (1) including direct and/or stationary or (preconditioned) Krylov subspace iterative methods. We refer to the survey of such methods and applications [1].

We concentrate here on the numerical behavior of two main segregated solution approaches. They are based on the transformation of the original saddle point system to the reduced one of a smaller dimension. When this system is solved (i.e. the component x or y), the remaining component is obtained by a back-substitution resulting to the another reduced system. Here we analyze iterative methods with short recurrences (e.g. one-step stationary methods, conjugate gradient method etc.) applied to the reduced systems which produce simultaneously the sequence of approximations x_k and y_k to the vectors x and y.

In the following sections, we present estimates on the attainable accuracy for the Schur complement method and the null-space projection method followed by illustration of these results on a numerical example. For details, we refer to the paper [7].

Our analysis is based on the work of Anne Greenbaum published in [2]. She exploited the well-known phenomenon that the updated residuals (e.g. in the conjugate gradients method) converge far below the machine precision (denoted by u here) for k large enough. When this is true, we can (approximately) bound the norm of the true residual by the norm of the difference between the true and updated one and thus to estimate the maximum attainable accuracy of certain iterative methods.

2. Schur complement reduction method

The Schur complement reduction method is based on the solution of the Schur complement system

$$B^T A^{-1} B y = B^T A^{-1} f \tag{2}$$

where $B^T A^{-1} B$ is the negative Schur complement of A in the (whole) saddle point matrix $\mathcal{A}$. This system is symmetric positive definite and it is solved by an iterative method which produces the sequence of approximations y_k to the solution component y. The corresponding approximation x_k to the vector x is then obtained by the substitution of y_k into the first equation of (1), i.e.

$$x_k = A^{-1}(f - B y_k).$$

The algorithms of the Schur complement reduction method are summarized in Figure 1. The iterative method to solve the Schur complement system (2) is characterized by Step 3, where we compute the direction vector $p_k^{(y)}$ and the step-length α_k (e.g. $p_k^{(y)} = r_k^{(y)}$, $\alpha_k = r_k^{(y)T} r_k^{(y)} / p_k^{(y)T} B^T A^{-1} B p_k^{(y)} = r_k^{(y)T} r_k^{(y)} / p_k^{(x)T} A p_k^{(x)}$ for the steepest descent method). We consider three different schemes for the computation of x_{k+1} in Step 7 which were used in many applications, e.g. in the context of the Uzawa method, two-level pressure correction algorithm etc., see [3, 4, 5, 6].

1. choose an initial guess y_0 (e.g. $y_0 = 0$), solve $A x_0 = f - B y_0$, compute $r_0^{(y)} = -B^T x_0$
2. for $k = 0, 1, 2, \ldots$
3. compute α_k and $p_k^{(y)}$
4. solve $A p_k^{(x)} = -B p_k^{(y)}$
5. update $y_{k+1} = y_k + \alpha_k p_k^{(y)}$
6. update $r_{k+1}^{(y)} = r_k^{(y)} - \alpha_k B^T p_k^{(x)}$
7. compute x_{k+1}:
   (a) update $x_{k+1} = x_k + \alpha_k p_k^{(x)}$
   (b) solve $A x_{k+1} = f - B y_{k+1}$
   (c) solve $A u_{k+1} = f - A x_k - B y_{k+1}$, update $x_{k+1} = x_k + u_k$
8. end

Figure 1: Schur complement reduction method: three schemes for computing x_{k+1}, (a) generic update, (b) direct substitution, (c) corrected direct substitution.

In the initial step (Step 1) of the algorithm and in Steps 4 and 7 of the iteration loop, we need to solve the system with the matrix block A. Thus the success of the Schur complement reduction method relies on the availability of a good approximation to the solution of the system with the matrix A. The analysis is based on the assumption that such a system is solved with the backward error equal to $\tau > 0$, i.e. the computed solution $\bar{v}$ of the system $Av = f$ is an exact solution of the perturbed system $(A + \Delta A)v = f$ with $\|\Delta A\| \leq \tau \|A\|$, where $\tau \kappa(A) \ll 1$ (ensuring the nonsingularity of the perturbed system). For the solution of a system with A, we use either the direct method (based on the Choleski factorization where $\tau = O(u)$) or the iterative method (the conjugate gradient method) with the stopping criterion based on the backward error.

2.1. Behavior in finite precision arithmetic

Here we recall results on the maximum attainable accuracy of the iterates $\overline{x}_k$ and $\overline{y}_k$ (we denote by bars the computed quantities). When the residual $\overline{r}_k^{(y)}$ converges far below the machine precision ($\overline{r}_k^{(y)} \simeq 0$), the

residual of the system (2) satisfies

$$\|B^T A^{-1} f - B^T A^{-1} B\overline{y}_k\| \lesssim \frac{O(\tau)\kappa(A)}{1-\tau\kappa(A)} \|A^{-1}\|\|B\|(\|f\| + \|B\|\overline{Y}_k), \tag{3}$$

where $\overline{Y}_k \equiv \max\{\|\overline{y}_i\| \,|\, i = 0, 1, \ldots, k\}$ is the maximum over the norms of the iterates $\overline{y}_i$ from the initial guess to the iterate in the step k. Here $\alpha \simeq \beta$ means that $|\alpha - \beta| \ll u$ and $\alpha \lesssim \beta$ means that $\alpha \leq \tilde{\beta}$ where $\tilde{\beta} \simeq \beta$. The estimate (3) shows, that the true Schur complement residual of (2) is proportional to the parameter τ (the backward error of the inner systems with A) and it also depends on $\overline{Y}_k$. The dependence on $\overline{Y}_k$, which can affect the attainable accuracy especially in the case of an irregular convergence of iterates for nonsymmetric problems, was also observed by Greenbaum in [2]. When the matrix $B^T A^{-1} B$ is symmetric positive definite, the convergence of iterates $\overline{y}_k$ is monotonic (at least in the exact arithmetic). When we choose $y_0 = 0$, we can usually replace $\overline{Y}_k$ by $\overline{y}$ which is the computed solution of (1).

In following paragraphs, we give bounds on the maximum attainable accuracy in terms of the norms of the residuals $f - A\overline{x}_k - B\overline{y}_k$ and $-B^T\overline{x}_k$.

Generic update $x_{k+1} = x_k + \alpha_k p_k^{(x)}$**:** The norms of the residuals of (1) satisfy

$$\begin{aligned} \|f - A\overline{x}_k - B\overline{y}_k\| &\leq O(\tau)(\|f\| + \|A\|\overline{X}_k + \|B\|\overline{Y}_k), \\ \|-B^T\overline{x}_k\| &\lesssim O(u)\|A^{-1}\|\|B\|(\|f\| + \|A\|\overline{X}_k + \|B\|\overline{Y}_k) \end{aligned} \tag{4a}$$

for large enough k, where $\overline{X}_k \equiv \max(\{\|\overline{x}_0\|, \|\overline{\alpha}_i\overline{p}_i^{(x)}\| \,|\, i = 0, 1, \ldots, k-1\})$. This update is "optimal", since it needs only one solution of the system with A per iteration step and produces the $\overline{x}$ which belongs to the null-space of B^T to the working precision. However, there is a growth of the norm of $f - A\overline{x}_k - B\overline{y}_k$ during the accumulation of the residuals of systems solved in Step 4 of the algorithm.

Direct substitution $x_{k+1} = A^{-1}(f - By_{k+1})$**:** The norms of the residuals of (1) satisfy

$$\begin{aligned} \|f - A\overline{x}_k - B\overline{y}_k\| &\leq O(\tau)(\|f\| + \|A\|\|\overline{x}_k\| + \|B\|\|\overline{y}_k\|), \\ \|-B^T\overline{x}_k\| &\lesssim O(\tau)\|A^{-1}\|\|B\|(\|f\| + \|A\|\overline{X}_k + \|B\|\overline{Y}_k) \end{aligned} \tag{4b}$$

for large enough k, where $\overline{X}_k \equiv \max(\{\|\overline{x}_0\|, \|\overline{x}_k\|, \|\overline{\alpha}_i\overline{p}_i^{(x)}\| \,|\, i = 0, 1, \ldots, k-1\})$. The use of the direct substitution scheme is suitable when the iterates $\overline{x}_k$ are not needed during the iteration loop and it is sufficient to compute it at the end.

Corrected direct substitution $x_{k+1} = x_k + A^{-1}(f - Ax_k - By_{k+1})$**:** The norms of the residuals of (1) satisfy

$$\begin{aligned} \|f - A\overline{x}_k - B\overline{y}_k\| &\lesssim O(u)(\|f\| + \|A\|\overline{X}_k + \|B\|\overline{Y}_k), \\ \|-B^T\overline{x}_k\| &\lesssim O(\tau)\|A^{-1}\|\|B\|(\|f\| + \|A\|\overline{X}_k + \|B\|\overline{Y}_k) \end{aligned} \tag{4c}$$

for large enough k, where $\overline{X}_k \equiv \max(\{\|\overline{x}_0\|, \|\overline{x}_k\|, \|\overline{\alpha}_i\overline{p}_i^{(x)}\| \,|\, i = 0, 1, \ldots, k-1\})$. The bound for the norm of $f - A\overline{x}_k - B\overline{y}_k$ holds starting with some k_0 and relies on the additional assumption that the iterates $\overline{y}_k$ begin to stagnate for $k \geq k_0$. It is similar to the residual update for the system $Ax_k = f - By_k$ and produces the accurate $\overline{x}_k$ for the computed $\overline{y}_k$ as indicated by the estimate.

Error estimates: Even though the residual estimates can differ depending on the formula for the computation of x_{k+1}, the norms of the errors $x - \overline{x}_k$ and $y - \overline{y}_k$ are of the same magnitude which is confirmed by the following bound:

$$\begin{aligned} \|x - \overline{x}_k\|_A &\leq \xi_1\|f - A\overline{x}_k - B\overline{y}_k\| + \xi_2\|B^T A^{-1} f - B^T A^{-1} B\overline{y}_k\|, \\ \|y - \overline{y}_k\|_{B^T A^{-1} B} &\leq \xi_2\|B^T A^{-1} f - B^T A^{-1} B\overline{y}_k\| \end{aligned}$$

with $\xi_1 \equiv \sigma_{min}^{-1}(A)$ and $\xi_2 \equiv \sigma_{min}^{-1}(B^T A^{-1} B)$. Both bounds have the residual of the Schur complement system on the right-hand sides, which is proportional to τ independently on the formula for x_{k+1}.

Backward error estimate for the generic update: We can also state the following backward error estimate. For large enough k (when $\overline{r}_k^{(y)} \simeq 0$), the iterates $\overline{x}_k$ and $\overline{y}_k$ are solutions of the perturbed system

$$\begin{pmatrix} A + \Delta A_k & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} \overline{x}_k \\ \overline{y}_k \end{pmatrix} \simeq \begin{pmatrix} f + \Delta f_k \\ \Delta g_k \end{pmatrix}, \tag{5a}$$

where

$$\|\Delta A_k\| \leq \gamma_k \|A\|, \ \gamma_k \equiv \tau \left(\|\overline{x}_0\| + \sum_{i=0}^{k-1} \|\overline{\alpha}_i \overline{p}_i^{(x)}\| \right) \left\| \overline{x}_0 + \sum_{i=0}^{k-1} \overline{\alpha}_i \overline{p}_i^{(x)} \right\|^{-1} \tag{5b}$$

and

$$\|\Delta f_k\| \leq \frac{O(u)\kappa(A)}{1 - \gamma_k \kappa(A)} (\|f\| + \|B\| \overline{Y}_k), \ \|\Delta g_k\| \leq \frac{O(u)\kappa(A)}{1 - \gamma_k \kappa(A)} \|A^{-1}\| \|B\| (\|f\| + \|B\| \overline{Y}_k). \tag{5c}$$

Equations (5) suggest that the computed solution is the solution of the perturbed system where the perturbation in the system matrix is proportional to τ and the perturbation in the right-hand side is proportional to u. The coefficient γ_k is generally greater or equal to τ, but starts to stagnate at some moderate level when the maximum attainable accuracy is achieved.

3. Null-space projection method

The second equation of (1) implies that the component x of the solution vector belongs to the null-space of the matrix B^T, i.e. x is determined by its components in $N(B^T)$ (N denotes the null-space of a matrix). Projecting the first equation of (1) onto $N(B^T)$ gives

$$(I - \Pi) A (I - \Pi) x = (I - \Pi) f, \tag{6}$$

where $\Pi = B^T (B^T B)^{-1} B$ is the orthogonal projector onto the range of B. This system is symmetric positive semidefinite and it can be solved by an iterative method which produces the sequence of iterates x_k approximating the solution component x. The corresponding iterate y_k can be found by solving the least squares problem

$$y_k = \arg \min_{w \in \mathbb{R}^m} \|f - A x_k - B w\|.$$

The algorithms of the null-space projection method are shown in Figure 2. In Step 3, we use the general two-term iterative method determined by the choice of the direction vector $p_k^{(x)}$ and the step-length α_k. Since we solve the projected system (6), the direction vectors must be restricted to $N(B^T)$. We consider four mathematically equivalent formulas for y_{k+1} in Step 7. The only update which is used in practice (and which we know about), is the generic update (see e.g. [8, 9]), but we include also schemes analogous to the case of the Schur complement reduction method.

In Steps 1, 4 and 7, the least squares problems (denoted by "$\approx$") with the matrix B need to be solved. Since we are not able to solve such systems exactly in practice, we suppose that the computed solution $\bar{w}$ of the least squares problem $Bw \approx g$ is the exact solution of the perturbed problem $(B + \Delta B)\bar{w} \approx g + \Delta g$ with $\max\{\|\Delta B\|/\|B\|, \|\Delta g\|/\|g\|\} \leq \tau$ where $\tau > 0$ is the backward error of the computed solution and it satisfies inequality $\tau \kappa(B) \ll 1$ ensuring that $B + \Delta B$ has a full rank. For the solution of $Bw \approx g$, we use the Householder QR decomposition giving $\tau = O(u)$ or the CGLS method with the stopping criterion based on the backward error.

3.1. Behavior in finite precision arithmetic

Here we will analyze the maximum attainable accuracy of the computed approximations $\overline{x}_k$ and $\overline{y}_k$. In the previous section, the updated residual $\overline{r}_k^{(y)}$ was an approximation to the residual of the Schur complement system. Now the projection of the updated resiual $(I - \Pi)\overline{r}_k^x$ is an approximation to the projected residual

1. choose an initial guess $x_0 \in N(B^T)$ (e.g. $x_0 = 0$), solve $By_0 \approx f - Ax_0$, compute $r_0^{(x)} = f - Ax_0 - By_0$
2. for $k = 0, 1, 2, \ldots$
3. compute α_k and $p_k^{(x)} \in N(B^T)$
4. solve $Bp_k^{(y)} \approx r_k^{(x)} - \alpha_k A p_k^{(x)}$
5. update $x_{k+1} = x_k + \alpha_k p_k^{(x)}$
6. update $r_{k+1}^{(x)} = r_k^{(x)} - \alpha_k A p_k^{(x)} - B p_k^{(y)}$
7. compute y_{k+1}:
   (a) update $y_{k+1} = y_k + p_k^{(y)}$
   (b) solve $By_{k+1} \approx f - Ax_{k+1}$
   (c) solve $Bv_k \approx -Ap_k^{(x)}$, update $y_{k+1} = y_k + v_k$
   (d) solve $Bv_k \approx f - Ax_{k+1} - By_k$, update $y_{k+1} = y_k + v_k$
8. end

Figure 2: Null-space projection method: four schemes for computing y_{k+1}, (a) generic update, (b) direct substitution, (c) updated solution, (d) corrected direct substitution.

$(I - \Pi)(f - A\overline{x}_k)$. As soon as the updated residual $\overline{r}_k^x$ converges far below the machine precision, we can give the following estimate:

$$\|(I - \Pi)(f - A\overline{x}_k)\| \lesssim \frac{O(u)\kappa(B)}{(1 - \tau\kappa(B))^2}(\|f\| + \|A\|\overline{X}_k), \tag{7}$$

where the quantity $\overline{X}_k$ is in this section defined as $\overline{X}_k \equiv \max\{\|\overline{x}_i\| \,|\, i = 0, 1, \ldots, k\}$. This bound is independent on the scheme of the computation of y_{k+1} and it shows that the first equation of (1) is satisfied accurately on the null-space of B^T to the working precision. However, this does not say anything about how the equation $Ax + By = f$ is satisfied and it actually depends on the formula for y_{k+1}. There is also another complication. Due to the restriction on the direction vectors to $N(B^T)$, the accuracy of the second equation of (1) depends on the algorithm used to solve the projected system (6). Since projections onto $N(B^T)$ are performed via the least squares problem with B, we can state (in the O-notation) the estimate

$$\| - B^T\overline{x}_k\| \leq O(\tau)\|B\|\overline{X}_k.$$

Now we show the estimates on the maximum attainable accuracy depending on the formula for y_{k+1}. Since the bound for the norm of $-B^T\overline{x}_k$ is common for all implementations, we give only estimates for the residual $f - A\overline{x}_k - B\overline{y}_k$.

Generic update $y_{k+1} = y_k + p_k^{(y)}$: The residual of the first equation of (1) satisfies

$$\|f - A\overline{x}_k - B\overline{y}_k\| \lesssim \frac{O(u)\kappa(B)}{(1 - \tau\kappa(B))^2}(\|f\| + \|A\|\overline{X}_k), \tag{8a}$$

for sufficiently large k. As in the case of the Schur complement reduction method, this update is optimal in the sense that only one least squares problem needs to be solved per iteration step and the first equation of (1) is satisfied accurately to the working precision.

Direct substitution $y_{k+1} = B^\dagger(f - Ax_{k+1})$**:** The residual of the first equation of (1) satisfies

$$\|f - A\overline{x}_k - B\overline{y}_k\| \lesssim \frac{O(\tau)\kappa(B)}{(1-\tau\kappa(B))^2}(\|f\| + \|A\|\overline{X}_k + \|B\|\overline{Y}_k), \tag{8b}$$

for sufficiently large k, where $\overline{Y}_k \equiv \max\{\|\overline{y}_i\| \,|\, i = 0, 1, \ldots, k\}$. This update could be suitable when we need not to have an approximation to y in every step k. However, the residual remains proportional to τ.

Updated solution $y_{k+1} = y_k + \alpha_k B^\dagger(-Ap_k^{(x)})$**:** The residual of the first equation of (1) satisfies

$$\|f - A\overline{x}_k - B\overline{y}_k\| \lesssim \frac{O(\tau)\kappa(B)}{(1-\tau\kappa(B))^2}(\|f\| + \|A\|\overline{X}_k + \|B\|\overline{Y}_k), \tag{8c}$$

for sufficiently large k, where $\overline{Y}_k \equiv \max\{\|\overline{y}_i\| \,|\, i = 0, 1, \ldots, k\}$. We have analyzed this update for the sake of completeness and it should not be prefered, since it needs the additional solution of the least squares problem with B and does not give good accuracy.

Corrected direct substitution $y_{k+1} = y_k + B^\dagger(f - Ax_{k+1} - By_k)$**:** The residual of the first equation of (1) satisfies

$$\|f - A\overline{x}_k - B\overline{y}_k\| \lesssim O(u)(\|f\| + \|A\|\overline{X}_k + \|B\|\overline{Y}_k), \tag{8d}$$

for sufficiently large k, where $\overline{Y}_k \equiv \max\{\|\overline{y}_i\| \,|\, i = 0, 1, \ldots, k\}$. This bound holds only when the iterates $\overline{x}_k$ start to stagnate from some index k_0.

4. Numerical experiments

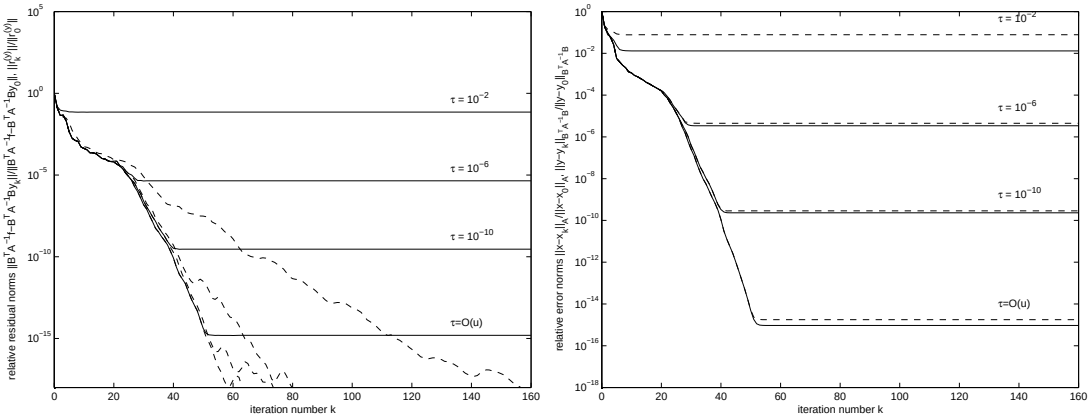


Figure 3: Schur complement reduction method: the relative norms of the residuals $B^TA^{-1}f - B^TA^{-1}B\overline{y}_k$ (solid lines) and the updated residual $\overline{r}_k^{(y)}$ (dashed lines) (on the left); the relative norms $\|x - \overline{x}_k\|_A/\|x - \overline{x}_0\|_A$ (solid lines) and $\|y - \overline{y}_k\|_{B^TA^{-1}B}/\|y - y_0\|_{B^TA^{-1}B}$ (dashed lines) (on the right).

We illustrate our theoretical results on the model problem taken from the finite element package IFISS for MATLAB. The testing problem is the Stokes problem

$$-\Delta\mathbf{u} + \nabla p = 0,\ \nabla\cdot\mathbf{u} = 0 \text{ in } \Omega,$$

where $\Omega = \{[x, y] \in \mathbb{R}^2 \,|\, -1 < x < 1,\ -1 < y < 1\}$. At the inflow boundary $x = -1$, the Dirichlet condition is set to $\mathbf{u} = (1-y^2, 0)$, at the boundary $y = -1$ and $y = 1$ we set $\mathbf{u} = (0, 0)$ and at the boundary $x = 1$, the Neumann condition $\partial_x u_x - p = \partial_x u_y = 0$ is prescribed representing the natural outflow. This problem represents the horizontal flow in the channel driven by the pressure difference between its two ends. The problem is discretized using "Q1–P0" finite elements with the discretization parameter $h = 1/8$ (thus $n = 578$ and $m = 289$). For more information, see e.g. the book by Elman et al [10].

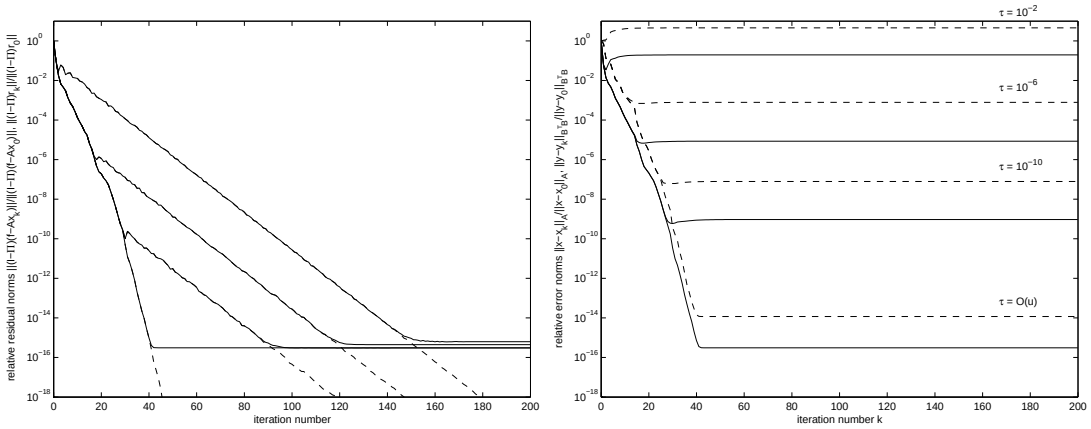


Figure 4: Null-space projection method: the relative norms of the projected residuals $(I - \Pi)(f - A\overline{x}_k)$ (solid lines) and the projected updated residual $(I - \Pi)\overline{r}_k^{(y)}$ (dashed lines) (on the left); the relative norms $\|x - \overline{x}_k\|_A / \|x - \overline{x}_0\|_A$ (solid lines) and $\|y - \overline{y}_k\|_{B^TB} / \|y - y_0\|_{B^TB}$ (dashed lines) (on the right).

τ	residual	Scheme		
		Generic update	Direct substitution	Corrected dir. sub.
$O(u)$	$\|f - A\overline{x}_{k_0} - B\overline{y}_{k_0}\|$	$6.1776{\cdot}10^{-15}$	$9.4674{\cdot}10^{-17}$	$4.0377{\cdot}10^{-17}$
	$\|-B^T\overline{x}_{k_0}\|$	$1.5092{\cdot}10^{-16}$	$1.3898{\cdot}10^{-15}$	$1.3970{\cdot}10^{-15}$
	$\|B^TA^{-1}f - B^TA^{-1}B\overline{y}_{k_0}\|$	$1.5495{\cdot}10^{-15}$	$1.5495{\cdot}10^{-15}$	$1.5495{\cdot}10^{-15}$
	$\|\overline{r}_{k_0}^{(y)}\|$	$6.9460{\cdot}10^{-42}$	$6.9460{\cdot}10^{-42}$	$6.9460{\cdot}10^{-42}$
10^{-2}	$\|f - A\overline{x}_{k_0} - B\overline{y}_{k_0}\|$	$2.1273{\cdot}10^{-1}$	$1.0223{\cdot}10^{-1}$	$1.6352{\cdot}10^{-16}$
	$\|-B^T\overline{x}_{k_0}\|$	$1.7414{\cdot}10^{-16}$	$5.6991{\cdot}10^{-2}$	$7.0881{\cdot}10^{-2}$
	$\|B^TA^{-1}f - B^TA^{-1}B\overline{y}_{k_0}\|$	$7.0457{\cdot}10^{-2}$	$7.0457{\cdot}10^{-2}$	$7.0457{\cdot}10^{-2}$
	$\|\overline{r}_{k_0}^{(y)}\|$	$4.3410{\cdot}10^{-19}$	$4.3410{\cdot}10^{-19}$	$4.3410{\cdot}10^{-19}$
10^{-6}	$\|f - A\overline{x}_{k_0} - B\overline{y}_{k_0}\|$	$1.6594{\cdot}10^{-5}$	$8.0717{\cdot}10^{-6}$	$3.7002{\cdot}10^{-17}$
	$\|-B^T\overline{x}_{k_0}\|$	$1.4235{\cdot}10^{-16}$	$4.9125{\cdot}10^{-6}$	$4.3974{\cdot}10^{-6}$
	$\|B^TA^{-1}f - B^TA^{-1}B\overline{y}_{k_0}\|$	$4.3974{\cdot}10^{-6}$	$4.3974{\cdot}10^{-6}$	$4.3974{\cdot}10^{-6}$
	$\|\overline{r}_{k_0}^{(y)}\|$	$8.4592{\cdot}10^{-32}$	$8.4592{\cdot}10^{-32}$	$8.4592{\cdot}10^{-32}$
10^{-10}	$\|f - A\overline{x}_{k_0} - B\overline{y}_{k_0}\|$	$1.2463{\cdot}10^{-9}$	$5.8143{\cdot}10^{-10}$	$3.0985{\cdot}10^{-17}$
	$\|-B^T\overline{x}_{k_0}\|$	$1.6542{\cdot}10^{-16}$	$3.2262{\cdot}10^{-10}$	$2.8892{\cdot}10^{-10}$
	$\|B^TA^{-1}f - B^TA^{-1}B\overline{y}_{k_0}\|$	$2.8892{\cdot}10^{-10}$	$2.8892{\cdot}10^{-10}$	$2.8892{\cdot}10^{-10}$
	$\|\overline{r}_{k_0}^{(y)}\|$	$1.9380{\cdot}10^{-37}$	$1.9380{\cdot}10^{-37}$	$1.9380{\cdot}10^{-37}$

Table 1: Schur complement reduction method: the norms of the residuals $f - A\overline{x}_{k_0} - B\overline{y}_{k_0}$, $-B^T\overline{x}_{k_0}$, $B^TA^{-1}f - B^TA^{-1}B\overline{y}_{k_0}$ and the updated residual $\overline{r}_{k_0}^{(y)}$ ($k_0 = 160$) for various values of τ and for all considered schemes (generic update, direct substitution, corrected direct substitution).

In Figure 3 we show the plots of the relative norms of the residuals of the Schur complement system (2) and the updated residuals $\overline{r}_k^{(y)}$ and the relative norms of the errors $x - \overline{x}_k$ and $y - \overline{y}_k$. These plots are taken from the scheme with the generic update, but the behavior of these norms is similar to all schemes. The plot on the left confirms the inequality (3). It shows that the residual $B^TA^{-1}f - B^TA^{-1}B\overline{y}_k$ is well approximated by the updated residual $\overline{r}_k^{(y)}$ until the level proportional to the parameter τ is reached. The similar behavior of the errors shown on the right plot illustrates the fact, that the errors of both components of the solution are proportional to the parameter τ independently on the formula for the computation of x_{k+1}.

Table 1 presents the values of residuals of the system (1), (2) and the updated residual in the step $k_0 = 160$. In this step, the updated residual is already far below the machine precision and therefore we can use the bounds (3) and (4) to estimate the attainable levels of the residuals $f - A\overline{x}_{k_0} - B\overline{y}_{k_0}$, $-B^T\overline{x}_{k_0}$ and $B^TA^{-1}f - B^TA^{-1}B\overline{y}_{k_0}$. The presented values of the residuals show a good agreement with their theoretical bounds.

In Figure 4 we show the plots of the relative norms of the residuals $(I - \Pi)(f - A\overline{x}_k)$ and the updated one

τ	residual	Scheme			
		Generic update	Direct substitution	Updated solution	Corrected dir. sub.
$O(u)$	$\|(I-\Pi)(f-A\overline{x}_{k_0})\|$	$3.0675{\cdot}10^{-16}$	$3.0675{\cdot}10^{-16}$	$3.0675{\cdot}10^{-16}$	$3.0675{\cdot}10^{-16}$
	$\|-B^T\overline{x}_{k_0}\|$	$1.0131{\cdot}10^{-17}$	$1.0131{\cdot}10^{-17}$	$1.0131{\cdot}10^{-17}$	$1.0131{\cdot}10^{-17}$
	$\|f-A\overline{x}_{k_0}-B\overline{y}_{k_0}\|$	$3.3632{\cdot}10^{-16}$	$6.2103{\cdot}10^{-16}$	$5.5196{\cdot}10^{-16}$	$3.1333{\cdot}10^{-16}$
	$\|\overline{r}^x_{k_0}\|$	$9.2584{\cdot}10^{-74}$	$9.2584{\cdot}10^{-74}$	$9.2584{\cdot}10^{-74}$	$9.2584{\cdot}10^{-74}$
10^{-2}	$\|(I-\Pi)(f-A\overline{x}_{k_0})\|$	$6.2076{\cdot}10^{-16}$	$6.2076{\cdot}10^{-16}$	$6.2076{\cdot}10^{-16}$	$6.2076{\cdot}10^{-16}$
	$\|-B^T\overline{x}_{k_0}\|$	$5.5589{\cdot}10^{-2}$	$5.5589{\cdot}10^{-2}$	$5.5589{\cdot}10^{-2}$	$5.5589{\cdot}10^{-2}$
	$\|f-A\overline{x}_{k_0}-B\overline{y}_{k_0}\|$	$1.1256{\cdot}10^{-15}$	$2.9618{\cdot}10^{-2}$	$1.6917{\cdot}10^{-2}$	$6.2477{\cdot}10^{-16}$
	$\|\overline{r}^x_{k_0}\|$	$1.3318{\cdot}10^{-20}$	$1.3318{\cdot}10^{-20}$	$1.3318{\cdot}10^{-20}$	$1.3318{\cdot}10^{-20}$
10^{-6}	$\|(I-\Pi)(f-A\overline{x}_{k_0})\|$	$4.4186{\cdot}10^{-16}$	$4.4186{\cdot}10^{-16}$	$4.4186{\cdot}10^{-16}$	$4.4186{\cdot}10^{-16}$
	$\|-B^T\overline{x}_{k_0}\|$	$1.5829{\cdot}10^{-6}$	$1.5829{\cdot}10^{-6}$	$1.5829{\cdot}10^{-6}$	$1.5829{\cdot}10^{-6}$
	$\|f-A\overline{x}_{k_0}-B\overline{y}_{k_0}\|$	$5.6027{\cdot}10^{-16}$	$2.5101{\cdot}10^{-6}$	$4.7047{\cdot}10^{-5}$	$4.4211{\cdot}10^{-16}$
	$\|\overline{r}^x_{k_0}\|$	$1.4050{\cdot}10^{-23}$	$1.4050{\cdot}10^{-23}$	$1.4050{\cdot}10^{-23}$	$1.4050{\cdot}10^{-23}$
10^{-10}	$\|(I-\Pi)(f-A\overline{x}_{k_0})\|$	$2.9949{\cdot}10^{-16}$	$2.9949{\cdot}10^{-16}$	$2.9949{\cdot}10^{-16}$	$2.9949{\cdot}10^{-16}$
	$\|-B^T\overline{x}_{k_0}\|$	$2.0535{\cdot}10^{-10}$	$2.0535{\cdot}10^{-10}$	$2.0535{\cdot}10^{-10}$	$2.0535{\cdot}10^{-10}$
	$\|f-A\overline{x}_{k_0}-B\overline{y}_{k_0}\|$	$3.9958{\cdot}10^{-16}$	$2.6784{\cdot}10^{-10}$	$1.6171{\cdot}10^{-9}$	$3.0651{\cdot}10^{-16}$
	$\|\overline{r}^x_{k_0}\|$	$2.2410{\cdot}10^{-26}$	$2.2410{\cdot}10^{-26}$	$2.2410{\cdot}10^{-26}$	$2.2410{\cdot}10^{-26}$

Table 2: Null space projection method: the norms of the residuals $(I-\Pi)(f-A\overline{x}_{k_0})$, $-B^T\overline{x}_{k_0}$, $f-A\overline{x}_{k_0}-B\overline{y}_{k_0}$ and the updated residual $\overline{r}^x_{k_0}$ ($k_0 = 200$) for various values of τ and for all considered schemes (generic update, direct substitution, updated solution, corrected direct substitution).

$(I-\Pi)\overline{r}^x_k$ and the relative norms of the errors $x-\overline{x}_k$ and $y-\overline{y}_k$. These plots are taken from the scheme with the generic update, but the behavior of these norms is also similar to all schemes. The plot on the left confirms the inequality 7. It shows that the projection $(I-\Pi)(f-A\overline{x}_k)$ of the residual $f-A\overline{x}_k-B\overline{y}_k$ is well approximated by the projected updated residual $(I-\Pi)\overline{r}^x_k$ until the level of the machine precision u is reached. Independently on the formula for y_{k+1}, the errors behave similarly and in all cases they ultimately stagnate on the level proportional to the parameter τ.

Table 2 presents the values of residuals of the system (1), the projected residual $(I-\Pi)(f-A\overline{x}_k)$ and the updated residual in the step $k_0 = 200$. In this step, the updated residual is already far below the machine precision and therefore we can use the bounds (7) and (8) to estimate the attainable levels of the residuals $f-A\overline{x}_{k_0}-B\overline{y}_{k_0}$, $-B^T\overline{x}_{k_0}$ and $B^TA^{-1}f-B^TA^{-1}B\overline{y}_{k_0}$. The presented values of the residuals show a good agreement with their theoretical bounds.

5. Conclusion

The main reason of this contribution is to analyze the maximum attainable accuracy of the various schemes of two main approaches used to solve the large-scale saddle point problems – the Schur complement reduction method and the null-space projection method. The attainable accuracy was measured in terms of both components of the residual of the system (1) and in terms of both components of the error $x-\overline{x}_k$ and $y-\overline{y}_k$. We have analyzed the influence of the rounding errors and the inexactness of the solution of the inner system (the system with the matrix A or the least squares problem with B). It was shown that the attainable accuracy measured in terms of the errors was similar for all considered schemes and it is proportional to the backward error in solving the inner systems. However, the behavior of the attainable accuracy in terms of the residuals of (1) was shown to be depenent on the chosen scheme of the computation of the secondary component of the solution. Some schemes lead ultimately to residuals that are on the level of the machine precision u; especially the cheapest generic updates give the approximations that satisfies either the first or the second equation of (1) to the working accuracy. The similar observation can be made in the case of the corrected direct substitution formulas.

References

[1] M. Benzi, G. H. Golub, J. Liesen, "Numerical solution of saddle point problems", *Acta Numerica*, vol. 14, pp. 1–137, 2005.

[2] A. Greenbaum, "Estimating the attainable accuracy of recursively computed residual methods", *SIAM J. Mat. Anal. Appl.*, vol. 18, pp. 535–551, 1997.

[3] H. C. Elman, G. H. Golub, "Inexact and preconditioned Uzawa algorithms for saddle point problems", *SIAM J. Numer. Anal.*, vol. 31, pp. 1645–1661, 1994.

[4] J. H. Bramble, J. E. Pasciak, A. T. Vassilev, "Analysis of the inexact Uzawa algorithm for saddle point problems", *SIAM J. Numer. Anal.*, vol. 34, pp. 1072–1092, 1997.

[5] A. Ramage, A. J. Wathen, "Iterative solution techniques for the Stokes and Navier-Stokes equations", *Int. J. for Numer. Meth. in Fluids*, vol. 19, pp. 67–83, 1994.

[6] J. Atanga, D. Silvester, "Iterative methods for stabilized mixed velocity-pressure finite elements", *Int. J. for Numer. Meth. in Fluids*, vol. 14, pp. 71–81, 1992.

[7] P. Jiránek, M. Rozložník, "Maximum attainable accuracy of inexact saddle point solvers", submitted, `http://flow.kmo.tul.cz/~paya/art/jr06.pdf`.

[8] N. I. M. Gould, M. E. Hribar, J. Nocedal, "On the solution of equality constrained quadratic programming problems arising in optimization", *SIAM J. Sci. Comp.*, vol. 23, pp. 1376–1395, 2001.

[9] D. Braess, P. Deuflhard, K. Lipnikov, "A subspace cascadic multigrid method for mortar elements", *Computing*, vol. 69, pp. 205–225, 2002.

[10] H. C. Elman, D. J. Silvester, A. J. Wathen, "Finite Elements and Fast Iterative Solvers with Applications in Incompressible Fluid Dynamics", Oxford University Press, 2005.

Downscaling in the context of data assimilation

Post-Graduate Student:
MGR. PAVEL JURUŠ

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2

182 07 Praha 8, CZ

Czech Technical University in Prague
Faculty of Transportation Sciences
Konviktská 20

110 00 Praha 1, CZ

jurus@cs.cas.cz

Supervisor:
DOC. ING. EMIL PELIKÁN, CSC.

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8

Czech Republic

emil@cs.cas.cz

Field of Study:
Engineering Informatics in Transportation and Telecommunications

This work was supported by the Grant Agency of the Academy of Sciences of the Czech Rep. (grant No. 1ET400300414, framework "Information Society") and by the Institutional research plan AV02 10300504 Computer Science for the Information Society "Models, Algorithms, Applications".

Abstract

Feasibility of common framework for the data assimilation and downscaling methods is investigated and discussed. Assimilation of measurements from ground level ozone stations with limited representativity to chemistry transport model is selected as an example and first results with simple decomposition to the local and background component are presented.

1. Introduction

Traditionally, data assimilation in environmental models is developed in the context of single scale. Resolution of a numerical model is given by the time discretization and grid spacing. Our typical objective is to find the 'analysis', which is as good estimate of reality as possible. In modern data assimilation techniques we usually combine information given by numerical model and by observations, as well as information about statistics of model and observational errors. Resulting analysis is defined on grid of numerical model and can be used, e.g., for further time integration or as initial and/or boundary condition for another coupled model.

Our efforts go beyond classical data assimilation – we don't want to have an estimate of model state as good as possible, but we also want to use information about model, observations and their respective errors for improved local predictions. This subject is often treated by downscaling methods or the model output statistics, but while some of these methods can utilize current observations, they don't use them to improve prediction of (coarse) numerical model and they don't use coarse model error statistics.

Local prediction is useful in number of cases. Here are some examples:

1. Wind direction and speed in location of wind farm or even specific wind turbine.
2. Prediction of road ice in given location.
3. Prediction of pollutant concentration in a specified place in an urban area.
4. Short-term prediction of wind field and of tracer dispersion in the vicinity of accidental release during dangerous goods transport.

All four cases are very useful applications of enviromental modelling that show several common features. We are interested in local properties of atmosphere and locality is of essential importance in these cases. While mesoscale NWP (Numerical Weather Prediction) models and/or CTM (Chemistry Transport Models) are the best sources of information available, they lack required space resolution. We need another model on top of mesoscale model(s) whose role is to downscale the prediction to the required location and resolution.

Models used for the downscaling vary depending on application and available data. The most widely investigated case of model output statistics is perhaps the first case above and models used in literature range from purely statistical regression and neural networks models to fine resolution CFD (Computer Fluid Dynamics) models.

Important aspect of aforesaid problems is that the finer the scale is, the more relevant is to predict the value *together* with its confidence interval or even to predict the whole probability distribution for non-gaussian quantities (e.g., road ice). Since the predicted quantity can be highly sensitive to the state of atmosphere (in atmospheric chemistry for example) and we lack the space averaging that helps in mesoscale models, the prediction errors might be very large in some cases. This means that sometimes the information about prediction error is even more important than the value itself.

Data assimilation framework gives us two basic benefits for downscaling. It gives us information about mesoscale state error statistics – this is especially straightforward for data assimilation based on ensemble methods. The simplest approach would be to perform downscaling for each ensemble member and regard the output as a sample from downscaled prediction. This approach is perhaps oversimplified because the resulting ensemble will lack information about error statistics of downscaling model, but there should be a remedy to this problem. The second benefit of data assimilation framework is the possibility to feed back the measurements into either mesoscale or downscaling model. This possibility is perhaps less pronounced for mesoscale NWP models where the analyses usually use so many various observations and quality is so high that assimilation of local observation might even deteriorate rather than improve the prediction. Usefulness of assimilation together with downscaling in mesoscale air quality model will be demonstrated in the next section and other possibilities of data assimilation also in downscaling model will be discussed in conclusion.

The first of listed cases is the example where a number of classical methods of downscaling is being used in practice. Assimilation of measurements (of wind or power production) back to the models is probably not important. Prediction of power production for a number of spatially distributed wind turbines can even help to cancel local wind turbulation and improve the forecast skill. Prediction is useful even without confidence intervals but probabilistic forecast would be certainly an enhancement.

Importance of probabilistic forecast is the main difference between the first and the second case. We are definitely interested in probability of road ice and a statistical model based on meteorological precursors could be appropriate solution. The design of a statistical model will depend on availibility of historical measurements and it could take into account the uncertainity in meteorological data.

The third case is an example when it makes sense to assimilate local measurements back into the mesoscale model. Exact method and the first results are selected as a demonstration of downscaling in data assimilation framework and will be main topic of the next section.

Final example is the most difficult among listed cases. There is usually not enough information to construct either deterministic or statistical downscaling model. There is no fine resolution model of terrain for CFD modelling and there is no history of model and observation values for statistical models – at least not at time of occurrence of an accident. Classical downscaling methods are therefore not applicable in this case. Design of model for local prediction in this case is a matter of further investigation.

2. Local observations in the air quality data assimilation

Data assimilating systems for tropospheric ozone modelling and prediction are currently being proposed and developed [1, 2]. The most common source of observations for air quality data assimilation are ground level stations that measure in situ concentrations of important pollutants. However, the spatial representativity of measurements differs greatly from station to station and can vary from order of 10^2m to 10^5m. Current practice is therefore only to assimilate the stations with the largest representativities – usually rural background stations, since only those stations have measurements aplicable on the scale of CTM. On the other hand most of stations are in urban areas where the air quality has the biggest impact and where mesoscale models are not suitable for modelling and forecasts of pollutant concentrations.

We have proposed and evaluated a data assimilating system for tropospheric ozone prediction with assimilation of rural background stations [3]. In this contribution we propose how to enhance system and how to allow it assimilate stations with smaller representativity and also how to improve local predictions of ground level ozone [4]. A test example of simple downscaling in data assimilation framework is also presented.

Our version of data assimilation is based on ensemble Kalman filter in square root formulation with localization and inhomogenous model error representation [5, 6]. Traditional discretized stochastic model of atmosphere evolution can be written as:

$$\mathbf{x}_k^t = \mathcal{M}(\mathbf{x}_{k-1}^t) + \nu_k \tag{1}$$

$$\mathbf{y}_k = \mathcal{H}(\mathbf{x}_k^t) + \varepsilon_k \tag{2}$$

where x_k^t is vector representing (inaccessible) discretized true state of atmosphere in timestep k on model grid, $\mathcal{M}$ is model of atmosphere used for time integration of state equation and ν_k is model error in timestep k. y_k is the vector of observation in timestep k, $\mathcal{H}$ is observation operator connecting observation and model state space and ε_k is its error in timestep k. ε_k is usually described as instrumental and representation error.

Observation operator in continuous space for a single observation $\mathbf{y}_i$ has usually the form [7]

$$y_i = \int_{\mathbb{R}^3} h(x)s(x, y_i)dx + \varepsilon(y_i) \tag{3}$$

where h is function connecting observation with physical space, and $s(x, y_i)$ aperture function depending on type of observation. For in situ observation is $s(x, y_i) = \delta(x - y_i)$ the Dirac delta function. For remote instruments (e.g. radars, lidars, satelites) implies $s(x, y_i)$ weighted area averaging over its support. For discretizated version it means that $\mathcal{H}$ is usually weighted average over values in some set of gridpoints – set of the nearest neighbours for in situ measurements or a larger set for remote observations.

We want to modify the equation (2) of stochastic model to employ a more general downscaling model instead of observation equation. The basic idea is to decompose ozone concentration to mesoscale component and local component. Decomposition is based on the assumption that part of ozone concentration have origin in long range transport of precursors and mesoscale weather conditions, the second part is added or substracted depending on local emmisions, deposition and other sub-mesoscale effects. The local component was estimated by very simple statistical model as 7-day moving average of the model update for the given hour of day – this reflects sometimes a strong diurnal cycle exhibited in ozone concentrations or a persistent bias. Modified stochastic model of atmosphere is now:

$$\mathbf{x}_k^t = \mathcal{M}(\mathbf{x}_{k-1}^t) + \nu_k \tag{4}$$

$$\mathbf{y}_k = \mathcal{H}(\mathbf{x}_k^t) + \mathbf{z}_k^t + \varepsilon_k \tag{5}$$

where the local component of ozone $\mathbf{z}_k^t$ is modelled as

$$\mathbf{z}_k^t = \frac{1}{7}\sum_{i=0}^{6} \mathbf{d}_{k-24i}^t + \eta_k \tag{6}$$

where $\mathbf{d}_k^t$ is model residual:

$$\mathbf{d}_k^t = \mathbf{y}_k - \mathcal{H}(\mathbf{x}_k^t) \tag{7}$$

We focus on testing the feasibility of the approach and do not try to find optimal solution for $\mathbf{x}$ and $\mathbf{z}$ simultaneously. We simplify solution by first estimating the local component $\mathbf{z}$ at the location of observations and then substituing our estimate into equation (5) and solving classical filtering problem. The solution gives us an estimate of mesoscale state. We want to estimate a vector of target variables $\mathbf{q}_k^a$ (as e.g. local ozone concentrations, probability of icing in specified point). For our testcase concerning ozone concetrations it is simply a sum of mesoscale analysis and local component $\mathbf{q}_k^a = \mathcal{H}(\mathbf{x}_k^a) + \mathbf{z}_k^a$, where superscript a denotes analysis. For linear observation operator and update equation for Kalman filter we can write:

$$\mathbf{q}_k^a = \mathcal{H}(\mathbf{x}_k^a) + \mathbf{z}_k^a = \mathcal{H}(\mathbf{x}_k^f) + \mathcal{H}(\mathbf{K}_k(\mathbf{y}_k - \mathcal{H}(\mathbf{x}_k^f) - \mathbf{z}_k^a)) + \mathbf{z}_k^a \tag{8}$$

Superscript f denotes forecast, K_k is Kalman gain matrix and $\mathbf{K}_k(\mathbf{y}_k - \mathcal{H}(\mathbf{x}_k^f))$ is analysis increment.

Forecast of mesoscale component is done by the NWP and CTM couple, and forecast of local component is done by persistent model, i.e. it repeats 24-hours old value. Index p denotes prediction horizon ($1 \leq p \leq 24$), and $\mathcal{M}^p$ means p successive applications of atmosphere model $\mathcal{M}$.

$$\mathbf{q}_{k+p}^f = \mathcal{H}(\mathbf{x}_{k+p}^f) + \mathbf{z}_{k+p}^f = \mathcal{H}(\mathcal{M}^p(\mathbf{x}_k^a)) + \mathbf{z}_{k+p-24}^a \tag{9}$$

Inspired by the equation (8) one could also alternatively try to predict analysis increment – in this case also by a simple persistent model:

$$\mathbf{q}_{k+p}^f = \mathcal{H}(\mathcal{M}^p(\mathbf{x}_k^a)) + \mathcal{H}(\mathbf{K}_{k+p-24}(\mathbf{y}_{k+p-24} - \mathcal{H}(\mathbf{x}_{k+p-24}^f) - \mathbf{z}_{k+p-24}^a)) + \mathbf{z}_{k+p-24}^a \tag{10}$$

3. Testcase and results

Testcase configuration was chosen to be as close as possible to operational system MEDARD [8] in the Institute of Computer Science in Prague, because one of the goals is to make data assimilation usable in operational mode. Selected NWP-CTM model couple was MM5 (PSU, NCAR) and CAMx (Environ) with NCEP FNL analyses as initial and boundary conditions. Model grid is identical to the outer domain of MEDARD system and has 80×65 cells with 27 km horizontal resolution, covering western and central Europe. CAMx model has 12 vertical levels and uses SAPRC-99 [9] chemistry mechanism.

Ozone epizode modelled took place in 24–29 June 2001. Two previous weeks were used for model calibration and spinup. We assimilated observations of O_3, NO, and NO_2 from the Airbase database [10]. 640 stations was selected for the assimilation – among them also mountain stations and urban background stations that are usually not assimilated due to representativity and/or bias problems.

We wanted to stay close to prospective operational mode and 24-hour tropospheric ozone prediction was made for the analysis at 5 p.m. Typical behaviour of local forecast compared to observation on station is illustrated in figures 1–5.

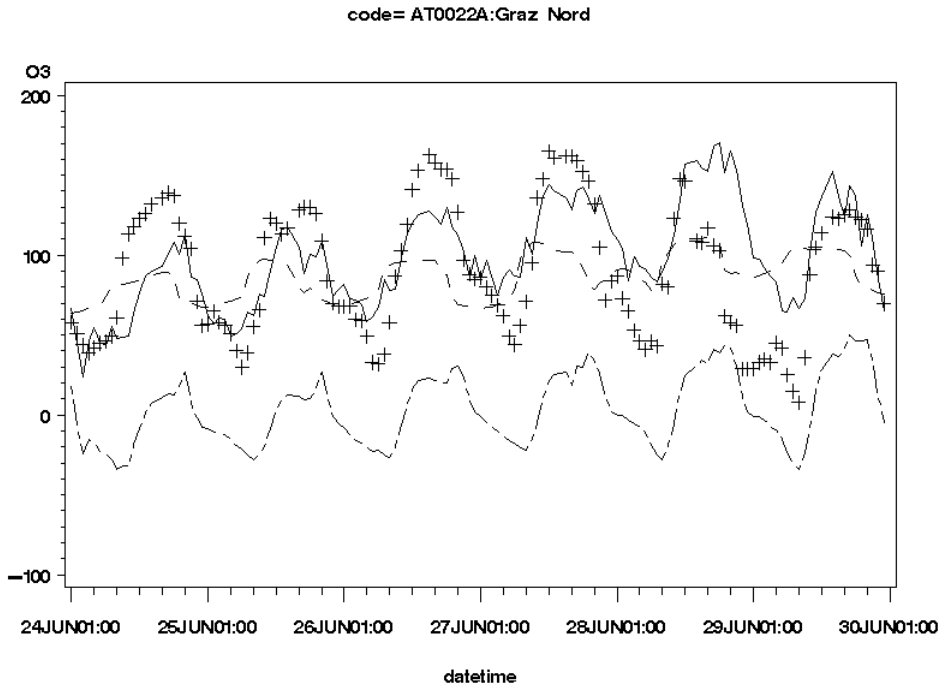


Figure 1: Observations and forecasts for selected station (Graz, Austria). Crosses: observations, dash-dot curve: estimate of local component, dash-curve: mesoscale model ($\mathcal{H}(\mathbf{x}^f)$), solid curve: target variable $\mathbf{q}_k^f$ containing term for analysis increment prediction (equation (10)).

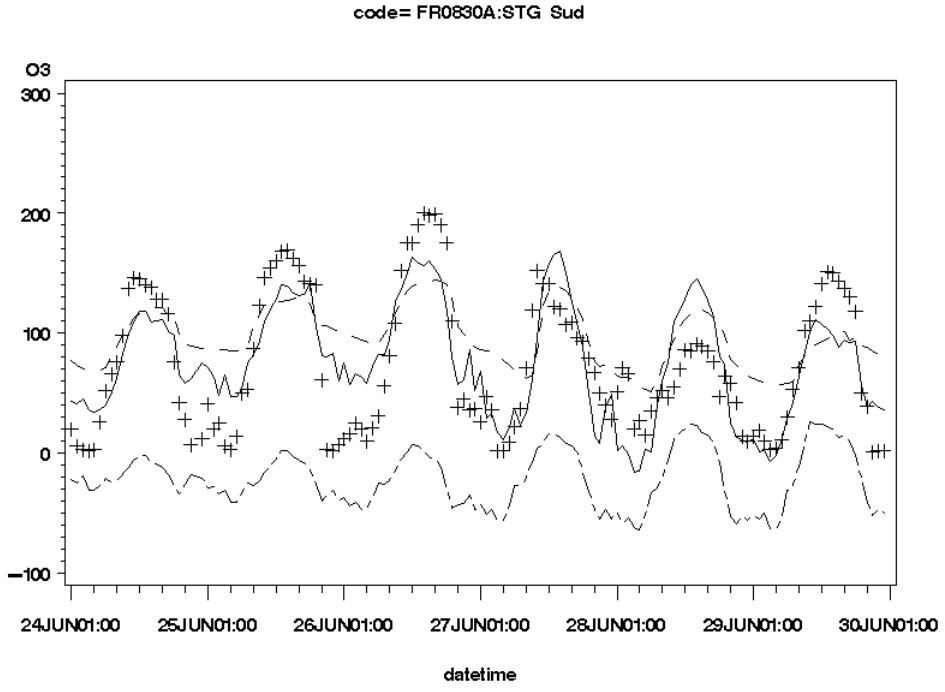


Figure 2: Observations and forecasts for selected station (Strasbourg, France). Crosses: observations, dash-dot curve: estimate of local component, dash-curve: mesoscale model ($\mathcal{H}(\mathbf{x}^f)$), solid curve: target variable $\mathbf{q}_k^f$ containing term for analysis increment prediction (equation (10)).

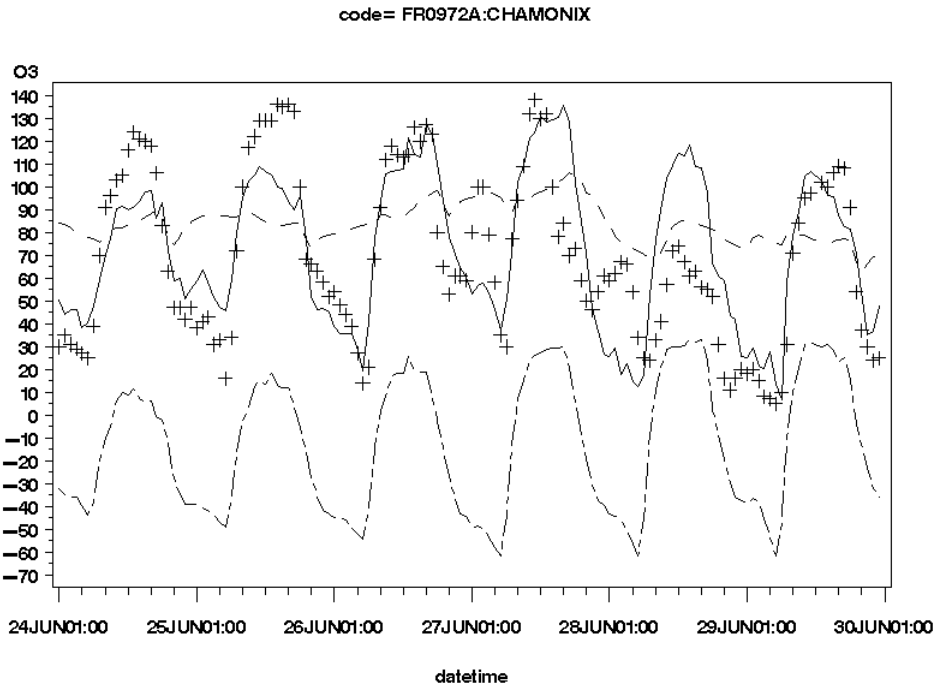


Figure 3: Observations and forecasts for selected station (Chamonix, France). Crosses: observations, dash-dot curve: estimate of local component, dash-curve: mesoscale model ($\mathcal{H}(\mathbf{x}^f)$), solid curve: target variable $\mathbf{q}_k^f$ containing term for analysis increment prediction (equation (10)).

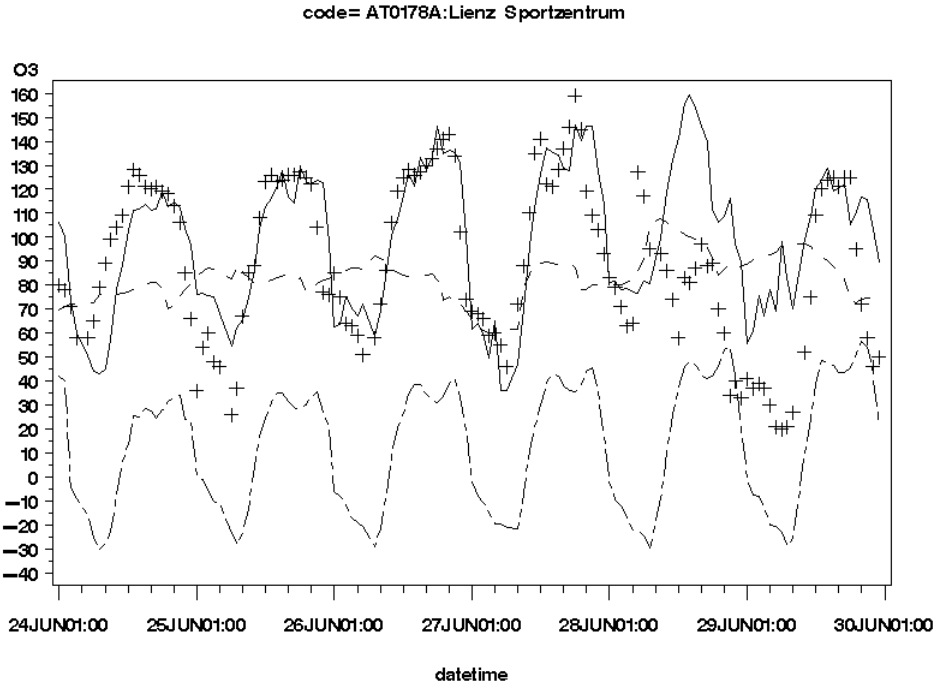


Figure 4: Observations and forecasts for selected station (Lienz, Austria). Crosses: observations, dash-dot curve: estimate of local component, dash-curve: mesoscale model ($\mathcal{H}(\mathbf{x}^f)$), solid curve: target variable $\mathbf{q}_k^f$ containing term for analysis increment prediction (equation (10)).

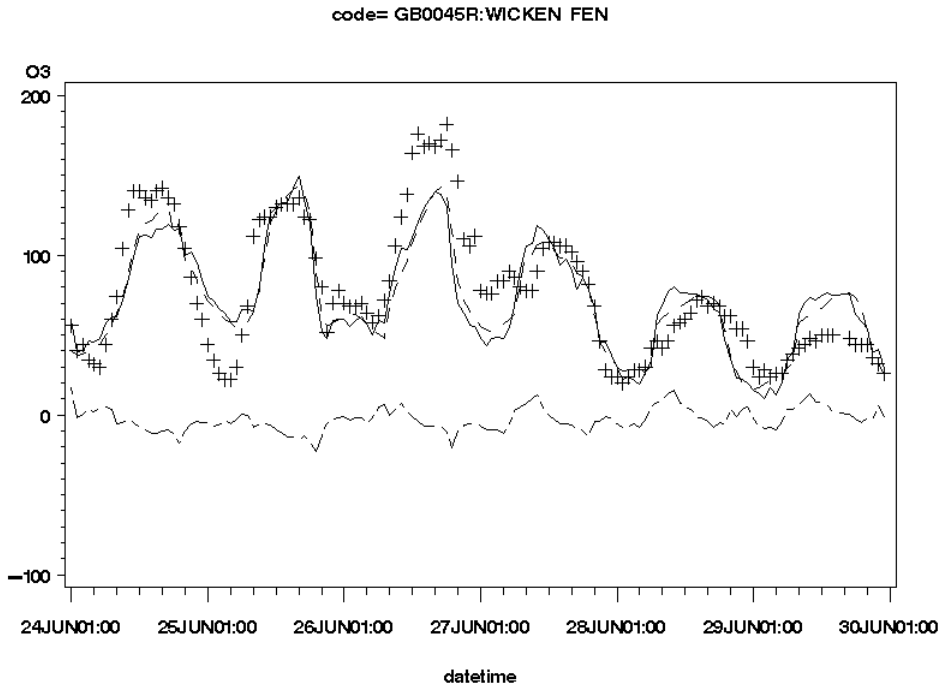


Figure 5: Observations and forecasts for selected station (Wicken Fen, UK). Crosses: observations, dash-dot curve: estimate of local component, dash-curve: mesoscale model ($\mathcal{H}(\mathbf{x}^f)$), solid curve: target variable $\mathbf{q}_k^f$ containing term for analysis increment prediction (equation (10)).

Figures 1 and 2 belong to urban stations with strong diurnal cycle. This cycle is not sufficiently described by mesoscale model. Model of local component rectifies this situation and local component increases with the beginning of the ozone episode. Figures 3 and 4 represent mountain stations with high orographic error and probably also inadequate traffic emmisions. Mesoscale model completely lacks the strong diurnal cycle that must be modelled by downscaling model. High persistence of simple moving average model causes worse performance in detection of the end of episode in the case of Lienz station. Figure 5 belongs to background rural station that is already accurately modelled by mesoscale model. Local component is negligible in this case.

First plot in figure 6 shows that forecast combining mesoscale model and downscaling model easily outperforms pure mesoscale model in prediction of ozone concentrations in location of measurement stations. Second boxplot shows that also classical statistical downscaling even with very simple statistical model brings some improvement over the pure mesoscale forecast. Improvement is seen in last two days of episode but not elsewhere. Last plot compares errors of classical statistical downscaling with statistical downscaling inside of data assimilation framework. The latter method is better in all days except the last one.

4. Conclusions and future work

Test case of data assimilation and downscaling in air quality model showed that for local forecasts even simple downscaling model is much better than mesoscale model alone. Downscaling within data assimilation framework exhibited better skill than model output statictics alone. The best results were obtained with predictor containing also the forecast of analysis increment. This could be attributed to the systematic part of air quality model bias. Combination of downscaling and mesoscale model also allowed us to use local observations that are usually unusable in traditional data assimilation due to large representativity errors. The testing period was quite short for a statistical verification but the first results are encouraging.

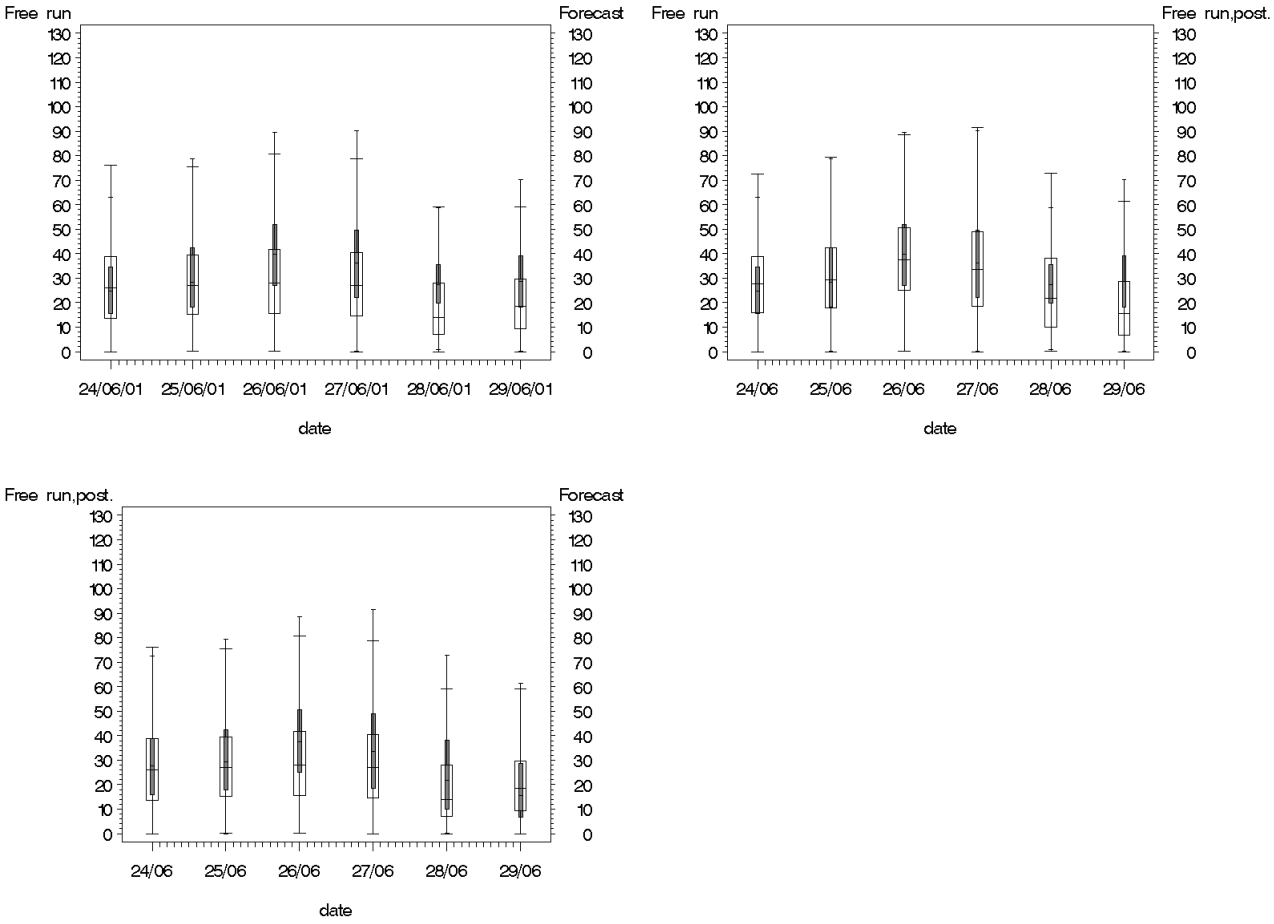


Figure 6: Boxplots of mean absolute errors of forecasts. Left side descriptions belongs to thin full boxes, right side belongs thick empty boxes. Center resp. bottom and top of boxes represent median resp. 1st and 3rd quartile of the error distribution. *Free run* is 1 day ahead forecast of mesoscale model started from analysed state. *Free run,post.* is 1 day ahead forecast of mesoscale model with separate downscaling from data assimilation. *Forecast* is 1 day ahead forecast of mesoscale model with the models of local component and analysis increment.

Our case demonstrated that combination of downscaling model, mesoscale model and data assimilation is beneficial even for very simple downscaling model. There are many possibilities for further development and improvement of presented methods. Downscaling model can be much more sofisticated. Explicit forecasts of the error covariances and confidence intervals which were omitted in our testcase is another topic that should be investigated. Analysis scheme can be improved to analyze simultaneously local and mesoscale component and provide better approximation of the best estimate in sense of the least squares or maximum likelihood. Careful modelling of error covariances is important in this case, since downscaling and mesoscale model can interact and compete for the explanation of residual variance.

There are further possibilities to investigate that are more specific for different cases. For example in case of nowcasting of accidental release (4th case in introduction), the model for downscaling can be a simple model accounting for terrain and landuse forcing of wind field. This model can adapt and refine its parameters with an increasing number of observations. Or the model for downscaling can be deterministic large-eddy simulation model. Data assimilation in traditional sense can be then performed in both mesoscale and large-eddy model.

References

[1] R. G. Hanea, G. J. M. Velders, A. Heemink, "Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport model", *J. Geophys. Res.*, vol. 109, D10302, doi:10.1029/2003JD004283, 2004.

[2] A. Segers, A. Heemink, M. Verlaan, M. van Loon, "A modified RRSQRT-filter for assimilating data in atmospheric chemistry models", *Environ. Mod. Software*, vol. 15, pp. 663–671, 2000.

[3] K. Eben, P. Juruš, J. Resler, M. Belda, E. Pelikán, B. C. Krüger, J. Keder, "An Ensemble Kalman filter for short term forecasting of tropospheric ozone concentrations", *Q. J. R. Meteorol. Soc.*, in print

[4] K. Eben, J. Resler, P. Juruš, "Accounting for representativity of ground-level air pollution monitoring stations in data assimilation systems", *Geophysical Research Abstracts* vol. 8, 10015, 2006.

[5] J. S. Whitaker, T. M. Hamill, "Ensemble data assimilation without perturbed observations", *Mon. Weather Rev.*, vol. 130, pp. 1913–1924, 2002.

[6] T. M. Hamill, J. S. Whitaker, "Accounting for the error due to unresolved scales in ensemble data assimilation: a comparison of different approaches", *Mon. Weather Rev.*, vol. 133, pp. 3132–3147, 2005.

[7] R. Daley, "Atmospheric Data Analysis", Cambridge University Press, New York, 1991.

[8] "The project MEDARD", Inst. of Computer Sci. Czech Acad. of Sciences, Prague. Website www.medard-online.cz

[9] W. P. L. Carter, "Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment", Final Report to California Air Resources Board Contract No. 92-329, and 95-308., 2000.

[10] "AirBase - the European air quality database", European Environment Agency, Website air-climate.eionet.europa.eu/databases/airbase

Onfram: nástroj pro lokalizaci biomedicínských ontologií

doktorand:
MGR. PETR KOLESA

Ústav informatiky Akademie věd ČR
Oddělení medicínské informatiky
Pod Vodárenskou věží 2
182 07 Praha 8

kolesa@euromise.cz

školitel:
ING. VOJTĚCH SVÁTEK, DR.

Katedra informačního a Iznalostního inženýrství
Vysoká škola ekonomická v Praze
Náměstí W. Churchilla 4
130 67 Praha 3

svatek@vse.cz

obor studia:
Biomedicínská informatika

Práce je podporována projektem 1ET200300413 a částečně i výzkumným záměrem AV0Z10300504.

Abstrakt

V článku popisujeme procedury a nástroj Onfram vyvinuté pro usnadnění a zrychlení tvorby lokalizovaných českých biomedicínských ontologií. Náš přístup je založen na vyhledávání konceptů v korpusu medicínských textů a jejich přiřazení protějškům v některé zavedené mezinárodní ontologii (tzv. *základové ontologii*). Takto vzniklá ontologie dvě hlavní výhody: je kompatibilní s některou z mezinárodních ontologií a potenciálně pokrývá všechny fráze používané v českém zdravotnictví.

Nástroj podporuje tvůrce ontologie tím, že se snaží automatizovat některé rutinní úkoly vyskytující se v průběhu tvorby ontologie. Nástroj se především snaží naučit, jak identifikovat koncepty v textu a jak je namapovat na základovou ontologii. Nástroj se učí postupně. Tak jak autor ontologie zpracovává další dokumenty, učící množina se zvětšuje a nástroj poskytuje lepší odhady. Po identifikaci konceptů, předkládá nástroj své návrhy uživateli. Ten buď přijme nebo opraví. Na základě této zpětné vazby nástroj upraví pravidla pro identifikací konceptů. Učení a úprava extrakčních pravidel probíhá pomocí metod zpracování přirozeného jazyka a nástroje na extrakci informací.

1. Úvod

Inteligentní medicínské systémy (IMS) pomáhají lékařům navrhnout a řídit léčbu specifickou pro jednotlivého pacienta a poskytují okamžité návrhy nebo upozornění o změnách v zdravotním stavu pacienta. Navíc napomáhají lékařům organizovat léčbu tak, aby byla v souladu s doporučenými postupy, ať se jedná o obecná doporučení vydávaná odbornými společnostmi nebo o technicko-organizační pravidla konkrétního zdravotnického zařízení. Takový přístup k léčbě vede k zlepšení poskytované péče na straně jedné a ke snížení výdajů za tuto péči na straně druhé. Z jiného úhlu pohledu IMS přebírají rutinní část lékařovi práce a lékař tedy může pracovat efektivněji.

Aby IMS mohly plnit tuto úlohu, musí být schopny pracovat s odpovídající medicínskou znalostí. Tato znalost musí být navíc poskytnuta ve formě, která je počítačově zpracovatelná. Dále musí vhodně popisovat oblast medicíny od elementárních skutečností až po velmi komplexní situace, a zároveň být pokud možno univerzální a znovupoužitelná, protože vytvoření takovéto znalostní reprezentace je nákladné a časově velmi náročné. Pro některé medicínské aplikace, jako například rozhodování v kardiologii, se zdá být nejlepší znalostní formalizací splňující výše vyjmenované *ontologie*.

K předním biomedicínským ontologiím patří *Foundation Model of Anatomy* (FMA) [1], *Systematized Nomenclature of Medicine* (SNOMED) [2] a *Unified Medical Language System* (UMLS) [3]. Všechny z těchto ontologií vznikly na území Severní Ameriky a soustředí se tedy především na podmínky a postupy obvyklé ve zdravotnictví v této části světa. Použití některé z těchto ontologií v aplikacích v českém zdravotnictví naráží kromě jazykové bariéry (všechny uvedené ontologie jsou primárně v angličtině a v různé míře přeloženy do francouzštiny, španělštiny, němčiny a několika dalších jazyků) i na rozdílnost zdravotnictví u nás a v Severní Americe. Příkladem takového rozdílu jsou používané léky: některé zde běžně používané léky nejsou dostupné v Severní Americe, a proto nejsou ani reprezentované v příslušné ontologii. Přímé

použití některé ze zavedených mezinárodních biomedicínských ontologií tedy není možné. Zároveň ale efektivně neexistuje žádná česká medicínská ontologie, a ani není žádná vyvíjena.

Bez vhodné biomedicínské ontologie ale není možné vyvíjet nebo nasazovat IMS do praxe v České republice. Tuto mezeru se snaží překlenout nástroj Onfram a metody, popsané v tomto článku. Onfram je nástroj určený k ulehčení vývoje biomedicínské ontologie vhodné pro vývoj IMS. Tvorba české biomedicínské ontologie je založená na tom, že v anglicky mluvících oblastech již existují IMS, které jako znalostní zdroj přímo využívají některou z výše jmenovaných ontologií.

Nejjednodušší postup, jak vytvořit ontologii, splňující výše uvedené požadavky je identifikovat v korpusu reprezentativních textů povrchové fráze [1] a namapovat je na koncepty v nějaké existující (*základové*) ontologii. Pokud takový koncept v základové ontologii neexistuje, je potřebný koncept vytvořen a autorem vložen do ontologie, včetně odpovídajících relací s ostatními koncepty.

2. Metody

Při vývoji nástroje jsme se zabývali otázkou, jak je možné zrychlit a usnadnit vývoj českých biomedicínských ontologií. Šlo nám hlavně o to, zautomatizovat některé části procesu vytváření ontologie, jako například identifikaci povrchových frází v textu a hledání odpovídajících protějšků v zavedené mezinárodní ontologii. Jsme tak schopni využit úsilí investované do vývoje dané ontologie. Tento způsob vytváření stále vyžaduje odborníka z oblasti tvorby ontologií, ale díky vyvinutému nástroji bude pracovat rychleji.

Česká ontologie vytvořená tímto způsobem není pouhým překladem základové ontologie, ale je odvozenou ontologii, která sice do značné míry přejímá strukturu základové ontologie, ale na rozdíl od ní dokáže popsat procesy v českém zdravotnictví. Odvozená ontologie navíc obsahuje pouze koncepty, které se v českém prostředí skutečně vyskytují. Na druhou stranu díky provázání odvozené a původní ontologie, bude jednoduché propojit systémy, založené na těchto ontologiích.

Ve dvou následujících oddílech stručně popíšeme základní metody, na kterých je Onfram založen. Ve třetí části bude stručně popsán vznik odvozené ontologie a ve čtvrté části rozebereme některé kroky tohoto procesu detailněji.

2.1. Extrakce informací

Obor nazývaný extrakce informací (information extraction, IE) se zabývá metodami automatického získávání informací z volného textu. Existuje mnoho metod a přístupů. V současnosti jeden z nejlepších algoritmů LP^2 [4], je implementován v nástroji AMILCARE [5], [6]. Tento nástroj jsme zvolili, protože podává dobrý výkon, a je vhodný pro interaktivní použití. AMILCARE navíc poskytuje Java API a je tedy snadno integrovatelný. AMILCARE využívá učení učitelem: na základě korpusu dopředu anotovaných textů (v každém dokumentu jsou vyznačeny informace, které chceme extrahovat) generuje systém extrakční pravidla. Na základě těchto pravidel jsou pak z neanotovaných textů extrahovány požadované informace.

AMILCARE není schopen sám o sobě provádět jazykovou analýzu textu. Toto není problém, pokud je text vhodně strukturován. Většina medicínských textů, jak se v průběhu vývoje ukázalo, však není dostatečně strukturovaná. Naštěstí je AMILCARE schopen využít dodatečných informací o textu, včetně jazykových.

2.2. Zpracování přirozeného jazyka

Zpracování přirozeného jazyka (Natural Language Processing, NLP) je obor, jehož zájmem je počítačové zpracování přirozeného jazyka takovým způsobem, jako by počítač do určité míry jazyku rozuměl. Porozumění jazyku počítačem je obecně velmi složitý problém, který neumíme dnes uspokojivě řešit.

[1] Povrchová fráze je část textu, která označuje nějaký konkrétní koncept. Různé povrchové fráze mohou označovat stejný koncept, například *glaukom* a *zelený oční zákal*. Na druhou stranu jedna povrchová fráze může v závislosti na kontextu označovat různé koncepty, *například teplota*: *fyzikální veličina* versus *symptom: zvýšená tělesná teplota*.

Nicméně existují nástroje, které pro praktické využití umožňují s uspokojivou přesností provádět morfologickou a syntaktickou analýzu textu.

Tyto nástroje jsou tak schopny poskytnout dodatečné informace o textu využitelné programem na extrakci informací, a tím zvýšit úspěšnost extrakce. K tomuto účelu využíváme český lemmatizer a morfologický analyzátor, který je součástí Pražského závislostního korpusu [7], [8]. Tento nástroj je schopen poskytnout řadu informací o zpracovávaném textu. V současnosti využíváme především informací o slovním druhu, pádu, čísle a slovním lemma (např. infinitiv u sloves, první pád jednotného čísla u podstatných jmen, zájmen, etc.). Lemmatizace je pro zpracování českých textů velmi důležitá, bez ní je prakticky nemožné ztotožnit různé tvary téhož slova a tedy identifikovat povrchové fráze označující stejný koncept.

2.3. Tvorba ontologie

Proces vytváření nové ontologie se sestává ze tří hlavních fází. V první fázi je zvolen způsob reprezentace znalostí. To zahrnuje rozhodnutí o způsobu reprezentace konceptů a vztahů mezi koncepty. Ve druhé fázi jsou identifikovány koncepty. Ve třetí fázi jsou nalezené koncepty začleňovány do ontologie – jsou definovány různé relace různých typů mezi koncepty. Ve skutečnosti je proces tvorby ontologie daleko komplikovanější – jednotlivé fáze se překrývají.

Pro úsporu času je při tvorbě ontologie racionální využít co největší množství práce, která již byla v oblasti tvorby ontologií vykonána. Proto se snažíme využít existujících mezinárodních ontologií jako základ pro vytvářenou českou ontologii. Nejprve tedy vybereme základovou ontologii tak, aby její struktura a definice typů a vztahů odpovídala zamýšlenému využití české ontologie. Tato část ontologie bude téměř beze změny převzata. Tento přístup je možný díky tomu, že, jak ukazují naše pozorování, existující mezinárodní biomedicínské ontologie ontologie jsou nevhodné k popisu jevů v českém zdravotnictví pouze proto, že neobsahují potřebné koncepty. Struktura typů a vztahů je vyhovující a proto může být převzata.

V druhé fázi je zpracováván korpus textů a jsou v něm vyhledávány povrchové fráze, které budou později mapovány na koncepty ontologie. Formálně je možné vytvořit korpus z libovolných textů, ovšem z hlediska úspěšnosti extrakce informací je vhodné, aby se jednalo o co nejvíce homogenní skupinu textů, ve kterých je informace odvoditelná spíše ze struktury a syntaxe než ze sémantiky. Příkladem takové množiny jsou například příbalové lékové informace. Při vhodně zvolených textech může být vyhledávání povrchových frází částečně zautomatizováno a tedy urychleno. Korpus je nástrojem zpracováván na pozadí tak, jak uživatel prochází další dokumenty v korpusu. Nástroj vyhledává povrchové fráze a uživatel jeho návrhy opravuje a doplňuje. Na základě této zpětné vazby systém zpřesňuje pravidla pro extrakci povrchových frází. Jak nástroj pro extrakci informací, tak nástroj pro zpracování přirozeného jazyka, byli vybrány tak, aby umožňovaly interaktivní reakci na rozhodnutí uživatele.

Ve třetí fázi je každá povrchová fráze identifikovaná v předchozí fázi přeložena do angličtiny a je hledán její protějšek v základové ontologii. Pokud je hledání úspěšné, je povrchová fráze přiřazena konceptu, pokud koncept nalezen není, je potřeba provést namapování ručně nebo v nové ontologii vytvořit koncept nový.

2.4. Extrakce povrchových frází a přiřazování konceptů

Při vytváření ontologie na zelené louce není k dispozici žádný počítačově zpracovatelný zdroj, kromě základové ontologie a jednoduchých slovníků jako například seznam diagnóz. Nejprve dojde k předzpracování korpusu textů: každý dokument je automaticky doplněn lingvistickými informacemi. V současnosti to jsou morfologické značky a slovní lemma. Ale obecně se může jednat o libovolnou informaci o syntaktické nebo sémantické rovině textu, která by mohla pomoct s identifikací povrchových frází. Příkladem jsou informace o struktuře věty.

V následující fázi uživatel prochází a zpracovává jednotlivé dokumenty. Po načtení dokumentu nástroj nejprve označí koncepty na základě nějaké pomocného slovníku, například seznamu léků [9]. Dále jsou aplikována extrakční pravidla generovaná na základě již zpracovaných dokumentů (při zpracování prvního dokumentu žádná taková pravidla neexistují). Povrchové fráze označené v předchozích dvou krocích jsou zobrazena uživateli k posouzení. Uživatel doporučení přijme, opraví nebo doplní. Na základě oprav

učiněných uživatelem, nástroj upraví extrakční pravidla.

V tomto okamžiku jsou všechny povrchové fráze v dokumentu identifikovány, ale mnoho z nich je ve tvaru složených jmenných frází, které musí být před dalším zpracováním rozloženy. Rozklad je v současné době prováděn heuristicky, algoritmem, který bere v úvahu podstatná jména, spojky, interpunkci a závorky ve složené jmenné frázi. Závorky obvykle obsahují příklad obecnějšího konceptu, který závorce předchází. Čárky a spojky oddělují jmenné fráze obsahující stejné podstatné jméno (případně rozvité), ale druhé a případné další výskyty tohoto podstatného jména jsou vynechány. Algoritmus se pokouší rozdělit složenou frázi a znovu zrekonstruovat jednoduché jmenné fráze, z nichž se skládá. Výsledné jednoduché jmenné fráze jsou zobrazeny uživateli jako návrhy, který má možnost je opravit.

Nakonec jsou povrchové fráze přiřazeny konceptům v základové ontologii. Proces přiřazení povrchových frází konceptu může být automatizován, pokud je daný koncept v základové ontologii obsažený. V našem testovacím případě je jako základové ontologie použito UMLS. Povrchové fráze (v češtině) jsou přeloženy do angličtiny pomocí česko-anglického slovníku slovnik.seznam.cz. Překladový modul zkouší překládat frázi jako celek i po jednotlivých slovech. Výsledkem překladu je tedy množina potenciálních výrazů. Tento přístup je možný díky tomu, všechny překlady jsou v dalším kroku testovány proti UMLS databázi a tak jsou vyloučeny všechny nesmyslné překlady.

Překlady povrchových frází jsou testovány pomocí aplikačního serveru UMLSKS [10]. Všechny koncepty, které se shodují s překladem povrchových frází, jsou pak zobrazeny uživateli, aby vybral ten správný. Uživatel má také možnost ručně zadat na který koncept se má povrchová fráze mapovat. Ručního mapování je vždy třeba, pokud se vhodný koncept v základové ontologii nenachází.

Onfram umožňuje vložit nový koncept na základě toho, že se v ontologii nachází *sourozenec* tohoto konceptu. Například antihistaminikum Flonidan, v UMLS neobsažené, bude mít většinu relací stejnou jako lék Zyrtec, tedy koncept v UMLS obsažený. Nástroj umožňuje projít relace, ve kterých se vyskytuje sourozenec (Zyrtec) nového konceptu (Flonidan) a pro nový koncept relevantní vztahy “zkopírovat”.

3. Výsledky

V průběhu vývoje byl nástroj Onfram testován na reálném problému: tvorbě české lékové ontologie na základě příbalových letáků. Nejprve byl vytvořen korpus 200 příbalových letáků, které byly náhodně vybrány z více než tří tisíc příbalových letáků léků dostupných v České republice. Úspěšnost nástroje byla měřena po zpracování poloviny dokumentů. Důvodem bylo, že extrakce povrchových frází je založena na učícím se algoritmu, jehož úspěšnost roste se zvětšující se trénovací množinou, tedy množinou již zpracovaných textů. Ostatní použité metody, jako je zpracování přirozeného jazyka nebo mapování povrchových frází na koncepty, nejsou v aktuální verzi Onframu ovlivněny zpětnou vazbou uživatele a proto jejich úspěšnost nezáleží na množství zpracovaných dokumentů.

Výsledky jsou uvedeny v tom pořadí, v jakém po sobě následují jednotlivé kroky, jsou v procentech a zaokrouhleny na celá čísla.

Nejprve proběhlo obohacení testu o jazykovou informaci. Morfologický analyzátor a lemmatizátor byl úspěšný v 96 procentech případů. To je velmi dobrý výsledek, vzhledem k tomu, že tento nástroj byl vytvořen na základě korpusu obecných textů, jako například novinové články, romány, etc.

Dále byly vyhledány v textu povrchové fráze. Nástroj pro extrakci informací dosáhl úplnosti (recall) 78 procent a přesnosti (precision) 93 procent. Úspěšnost rozkladu složených jmenných frází na jednoduché byla 64 procenta.

Při překladu byl použit on-line slovník slovnik.seznam.cz, který má dobré pokrytí v oblasti medicíny. Sedmdesát osm procent povrchových frází bylo přeloženo správně (v tom smyslu, že hledaný překlad byl v množině možných překladů, kterou vrátil překladový modul). Zhruba 11 procent povrchových frází bylo namapováno přímo prostřednictvím rozhraní UMLS Knowledge Source Server (UMLSKS), které je dos-

tupné na [10]. Toto rozhraní toleruje překlepy stejně tak unifikuje slova, jejichž pravopis kolísá. Tímto způsobem se dařilo najít protějšky slov latinského a řeckého původu, protože pravopis těchto slov se v češtině a angličtině liší jen lehce (například *hypokalemie* versus *hypokalemia*). Celkově bylo automaticky přeloženo a správně namapováno 75 procent povrchových frází. Uživatel pouze prováděl výběr v momentě, kdy bylo na výběr více možných mapování.

Pro zbylých 25 procent povrchových frází nebyl nástroj schopen navrhnout žádný mapování. Tato skupina frází se sestávala z frází, které mají v UMLS protějšek, ale nebyl nalezen (např. *člunkovitá deprese ST* versus *U-shaped ST depresion*) – celkově 17 procent všech frází; a frází, jejichž protějšek v UMLS nebyl schopen nalézt ani zkušený uživatel UMLS ve spolupráci s lékaři – 8 procent všech frází.

4. Diskuze

Nástroj Onfram je navržen především na tvorbu českých biomedicínských ontologií. Nicméně většina jeho komponent je jazykově neutrální a Onfram je navržen tak, že všechny jazykově závislé části, jako nástroje na zpracování přirozeného jazyka a překladový slovník, mohou být snadno nahrazeny. Stejně tak je možné Onfram použít na tvorbu některých ontologií i mimo oblast medicíny. Podmínkou je ale existence vhodné základové ontologie.

Největší slabinou aktuální verze nástroje Onfram je identifikace povrchových frází ve volném textu. V budoucnosti se chceme věnovat vylepšování úspěšnosti extrakce informací. Cestu vidíme jednak v poskytnutí dalších informací programu AMILCARE, jako je například větná struktura, jednak v kombinaci více nástrojů extrakcí informace, jako například AMILCARE a SVM^{light} [11].

Problémem extrakce informací a tvorby ontologií se zabývá několik projektů. Existuje zde tedy několik nástrojů, které poskytují tuto funkcionalitu. Nejznámějšími a nejvíce používanými jsou Protégé [12], nástroj na tvorbu ontologií, a GATE [13], obecný framework pro extrakci informací a zpracování přirozeného jazyka. Pro běžného uživatele zde však existue mezera – nástroj, který by komfortním způsobem spojoval oba nástroje. Proto jsem se rozhodli vyvinout vlastní nástroj, který by splňoval požadavky takového uživatele.

5. Závěr

V článku jsme popsali metody pro vytváření českých biomedicínských ontologiích. Dále jsme popsali nástroj Onfram, který tuto metody implementuje. Ukázali jsme, jakých výsledků nástroj dosahuje na reálném projektu – tvorbě české lékové ontologie. Onfram byl testován na korpusu 200 příbalových letáků. Mapování nalezených povrchových frází na koncepty ontologie je poměrně úspěšné (75 procent).

Náš přístup k vytváření českých biomedicínských ontologií je založen na využití struktury typů a vztahů některé z uznávaných mezinárodních biomedicínských ontologií, které jsou především v anglickém jazyce. Ukázali jsme, že není možné tyto ontologie přímo použít jednak kvůli jazykové bariéře a také kvůli chybějícím konceptům.

Nástroj Onfram zrychluje tvorbu odvozené české ontologie tím, že se pokouší automatizovat některé činnosti při tvorbě ontologie, jako například hledání povrchových frází ve volném textu nebo mapování povrchových frází na koncepty základové ontologie. Díky způsobu vzniku odvozené ontologie je tato do značné míry kompatibilní s původní – základovou – ontologií. Díky tomu, že koncepty jsou hledány v českých medicínských textech, odvozená ontologie obsahuje právě ty koncepty, které jsou v českém zdravotnictví používány.

Onfram se nesnaží plně automatizovat tvorbu ontologie, ani žádné z jeho částí. Cílem projektu je vyvinout nástroj, který pomůže uživateli tím, že redukuje rutinní práci, kterou by bylo jinak potřeba dělat ručně.

Literatura

[1] "The Foundational Model of Anatomy". *http://sig.biostr.washington.edu/projects/fm/index.html*.

[2] "Systematized Nomenclature of Medicine – Clinical Terms". *http://www.snomed.org/snomedct/*.

[3] "Unified Medical Language System." *http://www.nlm.nih.gov/research/umls/*.

[4] Ciravegna F. "(LP)2, an Adaptive Algorithm for Information Extraction from Web-related Texts." *In: Proceedings of the IJCAI-2001 Workshop on Adaptive Text Extraction and Mining*, 2001.

[5] "AMILCARE Homepage". *http://nlp.shef.ac.uk/amilcare/lp2.html.*

[6] Ciravegna F. "Adaptive Information Extraction from Text by Rule Induction and Generalization". *In Proceedings of IJCAI 2001*, 2001.

[7] "The Prague Dependency Treebank 1.0." *http://ufal.mff.cuni.cz/pdt/*.

[8] Hajič J. "Disambiguation of Rich Inflection - Computational Morphology of Czech". Karolinum, 2004.

[9] Státní ústav pro kontrolu léčiv "Databáze léčivých přípravků". *http://sukl.cz/cs02leciva/index.php*.

[10] "UMLS Knowledge Source Server". *http://umlsks.nlm.nih.gov/.*

[11] Joachims T. "Making large-Scale SVM Learning Practical. Advances in Kernel Methods - Support Vector Learning". *In B. Schölkopf and C. Burges and A. Smola (ed.)*, MIT-Press, 1999.

[12] "The Protégé Ontology Editor and Knowledge Acquisition System". *http://protege.stanford.edu/*.

[13] "General Architecture for Text Engineering". *http://gate.ac.uk/.*

Master-Slave PGA: Performance metrics

Post-Graduate Student:
ING. ZDENĚK KONFRŠT

Department of Cybernetics
Faculty of Electrical Engineering
Czech Technical University
Technická 2
16627 Praha 6

Czech Republic

konfrst@zikova.cvut.cz

Supervisor:
ING. MARCEL JIŘINA, DRSC.

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8

Czech Republic

marcel.jirina@cs.cas.cz

Field of Study:
Artificial intelligence and Biocybernetics

The research was supported by the CTU Grant No. 300008203–"Parallelization of Genetic algorithms".

Abstract

This work deals with a global model of PGA. A short overview of the model is introduced. Our research is concentrated on a master-slave algorithm (somewhere also called global). The complexity analysis of the algorithm is created and the processor optimality is derived from the analysis. Based on the optimality derivation, theoretical results with results from the real implementation are compared and the limitations of the algorithm are stated.

1. Introduction

In the literature, many classification of PGA models have been published. One of them divides PGAs into global, coarse-grained, fine-grained and hybrid types. More could be found at [1, 4]. We deal with the global model. It is also known as a master-slave model. There is only one population and evaluation of the population is spread to slaves. The main mechanism of a chosen PGA will be clarified later. It is generally well known that the algorithm is divided into two parts—a master and a slave part. We presuppose that there are more than one slave. Master process initializes variables and structures, creates the master task and writes to a process table. Then it creates slave processes and writes their tasks ids into the process table as well. After that it creates a first population and enters into while cycle. In the while cycle, the population is split according to the population size N and the number of processors p and sent to slave processes. When a slave finishes its evaluation, it sends back to master. An evaluated part is moved to a population array. When all slaves send their evaluations, the master starts a selection-crossover-mutation part.

2. Selection of PGA

We have several requirements on the PGA model. In first place, we want to keep the model as simple as possible, because the more difficult model we would implement, the more troubles we would get in our analysis. The models like coarse-grained, fine-grained and hybrid are powerful, but their optimization powers are a little bit frustrating for the analysis. As far as the model is concerned, we concentrate on a global PGA. There are many resources for it. The most important reason is to keep the run under simple GA rules and not to get confused by any indefiniteness. The second important reason is to focus on the study, the analysis and the design and not on the implementation. The implementation should confirm our theoretical predictions. The third reason is our worry of eventual problems, appearing during the implementation, which will come true in the end.

3. Selection of network and topology

We know from the introduction that the interconnection networks of PGA have crucial features. Optimalization is clear, to transfer maximum number of messages in the shortest time with minimum costs.

Usually, there are many contradictory requirements. The main requirements are: small and fixed degree, small diameter, symmetry, large bisection width, extendability and easy routing.

We chose a star network type. Let's the network is represented by a graph $G = (V, E)$, when $N = \sum_G V(G)$. It has the diameter $\phi(G) = 2$, the degree of a vertex $deg_G(u) = N - 1$, where $u \in V(G)$, the edge symetry, easy extendability.

4. Analysis

In the analysis, we employ performance metrics to evaluate the behaviour of the algorithm. For the solution, we take a classical master-slave architecture, which is induced by the chosen network. Not only is that an industrial standard, but it is easy to parallelize. This is also called a global PGA. By term "easy parallelize" we suppose easy to create a master and slave tasks, communications between them and keep them in a traditional GA area. During the run of a GA, procedures like selection, crossover, mutation and evaluation are employed. The last one–the evaluation of population is the most time consuming procedure, so therefore the evaluation in our approach is the task to divide to chunks and run in a parallel way. The master processor is responsible for selection, crossover, mutation and communication with slaves. The slave processor is responsible for evaluation and communication with the master.

In this section, we describe the main mechanism of the chosen PGA. We know that the program is divided into two parts–master and slave. We presuppose that there are more than one slave. **Master** processes initializes variables and structures, creates the master task and writes it to a process table. Then it creates of slave processes and writes their tasks ids into the process table as well. After that it creates a first population and enters into while cycle. In the cycle, the population is split according to N and p values and sent to slave processes. When a slave finishes it's evaluation, it sends the evaluation back to the master. An evaluated part is moved to a population array. When all slaves send their evaluations, the master starts a selection-crossover-mutation cycle. **Slave** receives the part of a population, evaluates it and returns it back to the master.

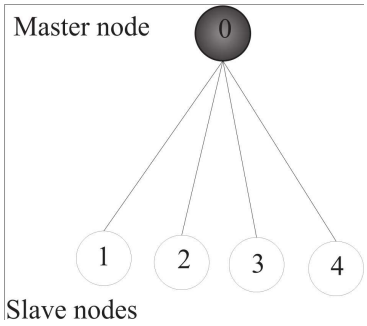


Figure 1: Topology for Master-Slave PGA. Master node is 0 and slave nodes are from 1 to 4.

5. Expected difficulty of the solution

When an estimation of behaviour of an algorithm is needed, we usually use performance metrics to evaluate it. We use X(N, p) to show that the task is of a size N running on p processors. We deal with the variables like the total parallel run time of an algorithm $T(N, p)$, the work $W(N, p)$, the cost $C(N, p)$, the speed up $S(N, p)$, the efficiency $E(N, p)$, the total overhead $H(N, p)$, processor optimality and other limitations. More about performance metrics are in [5].

6. Sequential GA

When we want to evaluate the population N and evaluation of one individual takes time T_o, we get total evaluation time T_{tot}. As we know GAs (PGAs) work in cycle, we have to keep in mind that we estimate only one cycle of an algorithm.

$$T_{tot}(N) = NT_o \tag{1}$$

Other sequential characteristics (W, C) can be easily extracted from the given equations (1).

7. Global Parallel GA

We are about to start with an estimation of the total parallel time $T_{tot}(N, p)$. Before that, we are going to explain some parameters: N is the total number of individuals, p is a number of processors, $p-1$ is a number of slave processors, T_w is the time when the master processor waits, T_e is the time for evaluation all population and T_c is the communication time between a master and a slave process.

$$T_w = T_e - (p-2)T_c \tag{2}$$

Let's explain the equation (2), why it holds. The master sends a portion of a population to a slave in time T_c. Then first slave starts to evaluate it's portion, while the master sends data to a second slave. The number of all slaves is $p-1$, but the first slave started evaluation after first T_c, that why we have $(p-2)T_c$. When all slaves are busy, the master is idle (T_w) until the first slave send message "ready".

If we presuppose that a task is equally divided and time for search T_e and individuals N on $p-1$ slave processors will be

$$T_e = \frac{NT_o}{p-1}, \tag{3}$$

when one position test takes T_o time. To sum up

$$T_w = \frac{NT_o}{p-1} - (p-2)T_c. \tag{4}$$

The total time T_{tot} will be:

$$T_{tot} = 2(p-1)T_c + T_w = \frac{NT_o}{p-1} + pT_c. \tag{5}$$

Precisely, we can write the total time of the parallel run as

$$T_{tot}(N, p) = T_{sel} + T_{cross} + T_{mut} + T_w + 2(p-1)T_c, \tag{6}$$

where T_{tot} is the total time of one cycle, T_{cross} is the time of a crossover process and T_{mut} is the time of a mutation. Adding T_{sel}, T_{cross} and T_{mut} to T_{scm}, where $T_{sel} > T_{cross} > T_{mut}$.

$$T_{tot}(N, p) = T_{scm} + \frac{NT_o}{p-1} + pT_c \tag{7}$$

The total time T_{tot} does not give us enough and complete information about behaviour the algorithm. So we will employ other characteristics to get better view in the analysis.

Let us start with the work $W(N,p)$ of the algorithm.

$$W(N,p) = T_c + 2T_c + 3T_c + \ldots + pT_c + (p-1)T_w + pT_c + (p-1)T_c + \ldots + T_c \tag{8}$$

$$W(N,p) = 2T_c(1 + \ldots + p) + (p-1)T_w \tag{9}$$

Then we can write

$$W(N,p) = p(p+1)T_c + NT_o - (p-1)(p-2)T_c. \tag{10}$$

After simplification so we can get

$$W(N,p) = NT_o + 2(2p-1)T_c. \tag{11}$$

$$W(N,p) \approx O(N+p)$$

Next characteristics is a total cost $C(N,p)$,

$$C(N,p) = NT_o + \frac{NT_o}{p-1} + p^2T_c. \tag{12}$$

Relation between $C(N,p)$ and $W(N,p)$ is $W(N,p) \leq C(N,p)$, when $p \geq 2$.

Is the algorithm cost optimal? No, it is not because of $O(N+p^2) \neq O(SU(N))$[1].

Next metric is a speedup $S(N,p)$,

$$S(N,p) = \frac{NT_o}{\frac{NT_o}{p-1} + pT_c} \approx \frac{p-1}{1+p^2k}, \tag{13}$$

where $k = \frac{T_c}{NT_o}$. The constant k is used in the next.

Efficiency $E(N,p)$ is

$$E(N,p) = \frac{NT_o}{NT_o + \frac{NT_o}{p-1} + p^2T_c}. \tag{14}$$

$E(N,p)$ from equation (14) is taken and simplified.

$$E(N,p) = \frac{NT_o}{NT_o + p^2T_c} \tag{15}$$

[1] In our case, $O(SU(N)) = O(SL(N)) = O(N)$.

We can change the equation above to the equation below by division of NT_o,

$$E(N,p) = \frac{1}{1+\frac{p^2T_c}{NT_o}} = \frac{1}{1+\frac{kp^2}{N}}, \tag{16}$$

we should minimize $\frac{kp^2}{N}$ to maximize the efficiency $E(N,p)$.

$$min(\frac{kp^2}{N}) \rightarrow max(E(N,p)) \tag{17}$$

Under conditions: $\frac{kp^2}{N} \rightarrow 1$, $E(N,p) \rightarrow \frac{1}{2}$, the size of population optimality is

$$N_{opt} = kp^2, \tag{18}$$

the size of processor optimality is

$$p_{opt} = \sqrt{\frac{N}{k}} \tag{19}$$

and total overhead $H(N,p)$ is

$$H(N,p) = p(\frac{NT_o}{p-1} + pT_c) - NT_o. \tag{20}$$

Let's make it simpler,

$$H(N,p) = \frac{NT_o}{p-1} + p^2T_c. \tag{21}$$

The optimal number of processors p_{opt} is received from the following:

$$\frac{\partial T_{tot}}{\partial p} = 0, \tag{22}$$

where T_{tot} is the total parallel run time from the equation (5) and p is the number of processors.

$$p_{opt} = \sqrt{\frac{NT_o}{T_c}} + 1 \tag{23}$$

8. Experiments

In the following, we describe carried out experiments and some achieved results. We also state fundamental computing and genetic parameters of the algorithm, because it makes the description more precise and worthy. The relevant parameters are enclosed in Table 1.

We evaluated p_{opt} and get to the result, which also corresponds with the measured values T_{totex}. If we interpret the result that the number of processor optimality is somewhere between 4 and 5, we might claim we got the right value. We used the one-max function as a simple test function. However, the test function is not a main issue, we suppose that the complexity of a test function would not change achieved results.

Computing parameters	
The number of workstations	7
Interconnection network	100Mbps Ethernet
The type of workstations	Sun Ultra Sparc
OS	Sun Solaris 8
Kernel version	SunOS 5.8
Message passing library	PVM 3.3
Genetic parameters	
Population sizes	$100 - 10000$
Population coding	binary
Crossover type	one point
Crossover rate	$0.8 - 1.0$
Mutation rates	$0 - 0.05$
Selection type	tournament
Tournament size	2

Table 1: The computing and genetic parameters of the algorithm employed in experiments.

9. Discussion

The disadvantages of MS topology are very well known. While all slave processes are computing, the master process is idle. While one slave communicates with the master, other slaves are waiting. During synchronization of generations, slaves are waiting.

In Figure 2, different total parallel run times per generation were measured. T_{tot} which stands as the theoretical total parallel run time, T_{totex} which stands for the experimental one. T_{tot2} has two times quicker under-laying communication environment and T_{tot10} ten times. The optimal number of processors p_{opt} could be calculated according to the equation (23). We reached the result as $p_{opt} = 4.16$ for the experimental system.

The model was implemented with a PVM library. The PVM [6] is a message-passing package that may use a network of heterogenous UNIX computers to behave like a single parallel computer. It could also be used on various parallel computers, so portability of software is possible. The PVM 3.3 library, the gcc compiler and a network of Sun SPARC workstations running Solaris OS were employed. The results are outcome of 100 independent runs. The tests were run in the Sun Sparc Laboratory at the Department of Computer Science and Engineering, where PVM 3.3 library and gcc compiler are installed. We acquired the knowledge and ideas from [2, 4, 6] to quicken the implementation phase. The results, presented here, confirmed our theoretical predictions and presupposed model limitations.

Our work covers the topic with $p \sim 10$. In this range, we understand that we loose a significant portion of a computing power by the idling master. To utilise the ranges $p \sim 100$, we need to have faster communication environments. The ranges $p \sim 1000$ and more are too large and we expect very expensive systems. The next important thing is that we did not investigate other models, their characteristics and behaviour. It was partly tackled in [2]. We made use of unfailing parameters from related works [1, 3, 4].

For $p \sim 10$, we could add a scheduler, which schedules jobs also to the master. When the master is idle, a small portion of evaluation could be processed. Some scheduling of slave jobs could also be investigated. As far as hardware characteristics are concerned, T_c and p are highly important. We run testing on a network of obsolete SPARCs. Nowadays, we can challenge new mighty technologies like Beowulf clusters supported by fast network environments (Gigabit Ethernets or Myrinets). Those will provide us with more free processors and a significant decrease of T_c to solve hard problems.

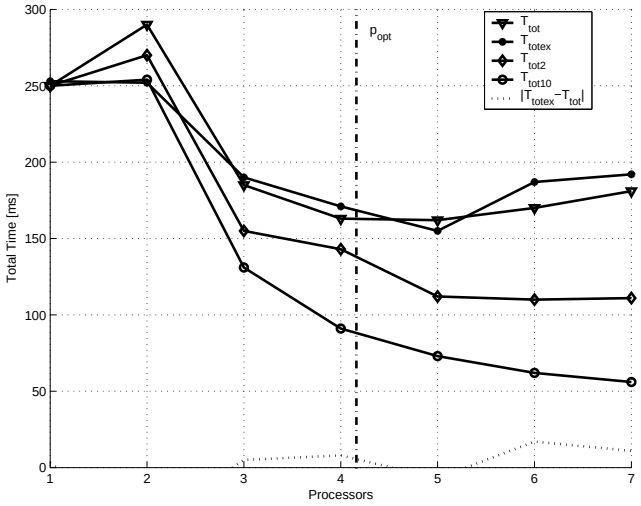


Figure 2: It shows theoretical predictions and experimental results of the number of used processors. The optimal number of processors is also presented. The most important graphs are T_{tot} and T_{totex}, because they show theoretical pattern and experimental reflection. Other curves (T_{tot2},T_{tot10}) are theoretical patterns for a faster network.

10. Conclusion

The complexity analysis of master-slave model was elaborated. Among others, we were focusing on $T(N,p)$, $W(N,p)$, $C(N,p)$, $S(N,p)$, $E(N,p)$, $H(N,p)$, the number of processors optimality and limitations of the parallel solution. The approximation characteristics of the model were introduced into space graphs. The limitation concerned worse scalability of the model, processors optimality and a demand on the exacting function evaluation. The derived equations (7, 23) work fine towards the experiments.

However, the ability of the model to utilize computing resources is low for $p \sim 10$, the global PGA brings a clear analysis, a comfortable implementation and a strict decrease of the total run time. As it was highlighted in the discussion, there are still many open questions for the further research in the area to run GAs faster and more efficiently.

References

[1] Konfršt Z.: *Parallel Genetic Algorithms*. Gerstner Laboratory Report 82/99, CTU FEE, Prague, 1999, p. 18.

[2] Cantú-Paz E.: *Implementing Fast and Realible Parallel Genetic Algorithms*. Handbook of Practical Genetic Algorithms, Vol. 3, 1999, p. 28.

[3] Cantú-Paz E., Goldberg, D. E.: Parallel genetic algorithms with distributed panmictic populations. IlliGAL Technical Report 99015, 1999, p. 8.

[4] Cantú-Paz E.: *A Survey of Parallel Genetic Algorithms*. IlliGAL Report No. 97003, 1997, p. 28.

[5] Tvrdík P.: *Parallel systems and algorithms*. Czech Technical Publishing House, 1994, pp. 167.

[6] Geist A., Beguelin A., Dongarra J., Jiang W., Manchek, R., Sunderam V.: *PVM 3 User's Guide and Reference Manual*. Oak Ridge National Laboratory, Tennessee, 1994, pp. 279.

Ontology-based Integration System

Post-Graduate Student:
ING. ZDEŇKA LINKOVÁ

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic

Department of mathematics
Faculty of nuclear science and physical engineering
ČVUT
Trojanova 13
120 00 Praha 2
Czech Republic

linkova@cs.cas.cz

Supervisor:
ING. JÚLIUS ŠTULLER, CSC.

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic

stuller@cs.cas.cz

Field of Study:
Mathematical Engineering

This work was supported by the project 1ET100300419 of the Program Information Society (of the Thematic Program II of the National Research Program of the Czech Republic) “Intelligent Models, Algorithms, Methods and Tools for the Semantic Web Realization” and by the Institutional Research Plan AV0Z10300504 “Computer Science for the Information Society: Models, Algorithms, Application”.

Abstract

Integration has been an acknowledged problem for a long time. With the aim at combining data from different sources, data integration usually provides a unified global view over these data. A crucial part of the task is the establishment of the connection between the global view and the local sources. Two basic approaches have been proposed for this purpose: Global As View (GAV) and Local As View (LAV). With the Semantic Web and its data description means, there is also another possibility - to employ ontologies for the relationship description in an integration system.

1. Introduction

Today’s world is a world of information. The expansion of World Wide Web has brought a number of information sources. However, at the same time, a number of different formats, data heterogeneity, and not yet efficient machine processing of web sources cause many problems. One of them is the reappeared problem of data integration.

Data integration is the task of combining data residing at different sources and enabling the user to process these data as one whole. Data integration has been an acknowledged data processing problem for a long time. Although there have been some projects on integration of data within particular areas, there is no universal tool for general data integration.

In general, data integration can be pursued in different layers. It is possible to consider only data, or consider also metadata (e.g. schemas). With greater data amount, the integration approach is rather non-materialized than materialized. The integration result brings virtual view over data sources that do not store any data. Therefor, the establishment of the connection to original data sources is crucial. To consider the data schemas is essential. There are some basic approaches to the design a non-materialized integration system, each with its advantages and disadvantages. The proposed approach brings an idea from the Semantic Web - a semantic extension of the current World Wide Web.

The paper is organized as follows. Section 2 describes the data integration task and basic approaches. Section 3 introduces the vision of the Semantic Web and one of its principle layers - ontologies. Section 4 presents an ontological approach to data integration. Finally, Section 5 summarizes the paper.

2. Data Integration

In data integration, the goal is to synthesize data from different data sources into one integrated data source. A user willing to process the data uses the integrated source and is freed from the knowledge where the data are and how the data are structured in the respective sources.

The integrated data source can be materialized, i. e. a new data source is created and it physically stores data, or it can be virtual, i. e. a virtual view is defined and the data remain in the sources. In materialized data integration approach, a copy of the data is made. So, with respect to actualization requirements, it is suitable for more or less stable data. Virtual data integration approach provide an interface to autonomous data sources, it can be used also for large amount of data with relatively frequently changing content. In a connection with the World Wide Web data, this approach suggests itself. It is also possible to combine both approaches. An example is an integration system that provides a virtual integrated view, but it also materializes some data in a cache. The cache is usually used for frequently accessed data.

A commonly used system architecture in virtual view approach [1] to data integration is depicted in Figure 1. A base of the system is a set of data sources to integrate. The higher layer is a set of components called wrappers. Each wrapper belongs to one local data source and it plays a role of a connector between the local source background (it means a specific model, a specific language etc. for the source) and the global one. The pure integration part of the system is presented in hierarchical layers of mediator components. A mediator can obtain information from components below it and can provide information to components above it. In general, an integration system can contain an arbitrary complex architecture.

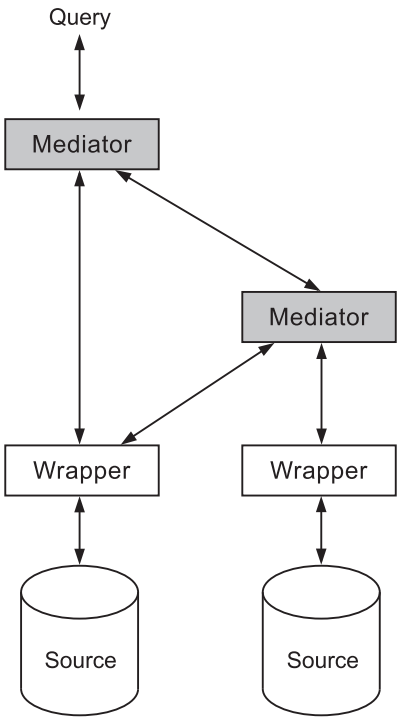


Figure 1: A mediation integration architecture

Each mediator in a hierarchy can be seen as a virtual view. These views are then used in query evaluation. A user of the integration system poses his query to a global view using a global schema. Using mediation integration, the query is reformulated and decomposed to refer to the data sources and the queries are also executed over the sources. Then obtained information is composed and the answer is given back to the user.

The main components of a data integration system are the sources with their local schemas, the global virtual view with the global schema, and the mediated system that expresses the correspondence between the global source and the local sources. So it follows that a data integration system I is a triple

$$I = (G, L, M),$$

where G is a global schema, L is a set of local schemas and M is a mediation system.

A possible way how to describe the mediation system is to use mappings. Mapping is a set of assertions that establish the connection between the element of the global schema and those of the local schema. The composition of mapping is an essential task. It plays a crucial role in query resolving - another important process of a data integration system. Two basic approaches [2, 3] have been used in order to specify the mapping. The *Global As View* (GAV) approach consists in defining the global schema as a set of views over the local schemas, while the *Local As View* (LAV) approach consists in defining the local sources as a set of views made on the global schema.

Because the GAV is based on the idea that the content of each element of the global schema should be characterized in terms of a view over the sources, this mapping tells the system how to retrieve the data. The GAV favors the system in carrying out query processing - it gives direct information on how query answering may be performed. Some GAV data integration systems do not allow integrity constraints in the global schema. Under these assumptions, query processing can be based on a simple unfolding strategy: every element of the global schema is substituted with the corresponding query over the sources. When global schema allows integrity constraints, the query processing in GAV becomes more complex - integrity constraints here can in principle be used in order to overcome incompleteness of data in the sources. In GAV query processing can look easy. However, this idea is effective when a set of sources is stable. The addition of a new source and extending the system can be difficult. The new source may have an impact on the definition of various elements of the global schema. So it can force the system designer to redesign the schema, and so to reconsider all the sources.

The LAV approach is based on the idea that the content of each source can be characterized in terms of a view over the global schema. Processing queries in LAV is a difficult task. The only knowledge we have about the data in the global schema is through the views representing the sources, and such views provide only partial information about the data. The mapping specifies the role of each source relation with respect to the global schema. It is not immediate to infer how to use the sources in order to answer queries. The LAV favors the system in the extensibility - addition of sources simply means enriching the mapping with definition of a new view over the global schema, without other changes.

To compensate the insufficiency of the LAV and GAV approaches, also their combinations have been proposed. The *Global Local As View* (GLAV) approach [4] establishes the relationships between the global schema and the sources by making both of LAV and GAV mappings and allows flexible schema definitions independent of the particular details of the sources.

3. The Semantic Web

The Semantic Web [5, 6] is intended as an extension of today's World Wide Web. It should consist of machine readable and efficiently processable data. The basis is addition of data semantics - data description will be stored together with data themselves. The full realization of the Semantic Web belongs still to the future; however, many tools, languages, theories etc. have been developed and several also implemented. The Semantic Web is based on several standards, which are defined by W3C (WWW Consortium) [7].

An important requirement for effectively machine processable data is data structuring. On the web, the main structuring method is using tags, which are parts of text containing information about the role of the text. Nowadays, the metalanguage XML (eXtensible Markup Language) [8] is used for making web document structure. It provides syntax for machine readable data. But only XML is not enough to describe data. The technique to specify the meaning of information is RDF (Resource Description Framework) [9]. It is a

basic tool of web sources metadata addition. RDF data model gives an abstract conceptual framework for metadata definition and usage. It uses XML syntax (RDF/XML) for encoding. Additionally, there is also an extension of RDF called RDF Schema [10] that is useful for class definition and class hierarchy description.

An instrument for definition of terms used either in data or in metadata are ontologies [11]. The term ontology has been used in many ways and across different communities. A popular definition of the term ontology in computer science is: an ontology is a formal, explicit specification of a conceptualization. A conceptualization refers to an abstract model of some phenomenon in the world. However, a conceptualization is never universally valid. Ontologies have been set out to overcome the problem of implicit and hidden knowledge by making the conceptualization explicit. Ontologies aim at modeling and structuring domain knowledge. It may take a variety of forms, but it will necessarily include a vocabulary of terms and some specification of their meaning. In the context of web technologies, ontology is a file or a document that contains formal definitions of terms and term relations. The Semantic Web technique for definition of ontologies is the OWL (Ontology Web Language) language [12].

4. Ontology-based mediation integration

In an ontology-based integration approach described in this paper, a conception of a virtual integration form is adopted. A global source will be also non-materialized and for the establishment of a connection to the data sources some kind of mapping will be applied. However, instead of using mapping rules as assertions for global and local schemas elements, a more complex structure covering all mapping will be employed. This approach exploits the idea that on the Semantic Web every peace of information has got defined its meaning and supposes availability of ontologies as a means for defining the concept of the data. The integration task is transformed to the building of an ontology for the integration system. This ontology from its principle should cover ontologies of all data used in the system and mappings that are in general seen as definitions of relationships between data.

Suppose, there are two data sources S_1 and S_2. Each source schema is described by an ontology: an ontology referring to the local source S_1 is O_{S_1}, an ontology of the source S_2 is O_{S_2}. The global integrated view the integration system should provide has an ontology O_G. The integration system, in Section 2 formalized as a triple $I = (G, L, M)$, has in this case representation

$$I = (O_G, \{O_{S_1}, O_{S_2}\}, O_I),$$

where O_I is an ontology of the integration system.

Ontology O_I is used to describe the mapping between elements of the global view and the local sources. O_I is also an ontology of all concepts used in the integration system I. So it follows that for ontologies of local sources is valid:

$$O_{S_1} \subseteq O_I$$
$$O_{S_2} \subseteq O_I$$

While ontologies O_{S_1} and O_{S_2} are given with the sources, O_G and O_I need to be determined. Description of O_G is relatively independent on the sources. O_G contains definition of concepts accessible directly via the global view. It is a matter of a designer who decides what will be accessible via the integration system and in what form.

Establishment of O_I is a crucial step. However, it is not an easy task. Covering O_{S_2}, O_{S_2}, O_G and their relationships, O_I is the result of task called merging ontologies. Ontology merging is studied e.g. in [13] and [14]. As in schema integration in other approaches, some conflicts [15] that have to be solved can arise. Conflicts can be of various types [16], for example terms conflicts, schema discrepancies, raw data and metadata conflicts etc. Regarding used terms, synonym and homonyms conflicts can arise. (Synonyms are different words with similar or identical meanings. Homonyms are words that are spelled the same but

with a different meaning.) In the ontology world, it is not difficult to deal with synonyms or homonyms, because there are means how to express relationship between terms. In an ontology, each term has a unique reference. Although, there can be two terms in two ontologies named in the same way, they are uniquely distinguishable, because of the context - the ontology where they are defined. This is for instance in XML syntax solved by namespaces. Within ontologies it is also possible to state that two terms are equivalent and describe the synonymic relationship, and by this to enable process it in a right way.

Example 1 There are two sources to integrate. Source 1 stores satellite images taken from a satellite Ikonos and its ontology O_1 describes only one class named `Ikonos_images` with properties `date_`,`the_geom` etc. Source 2 stores satellite images from a satellite Spot, its ontology O_2 contain class `Spot_images` with properties `date_acqui`, `the_geom` etc. Since the integration system should provide satellite images coming from different data sources, global ontology O_G contains class named `satellite_images` with properties `date` (date when the image was taken), `geom` (a geometry of the photographed region), etc. To obtain ontology of the system, O_1, O_2, O_G, and knowledge about relationships among particular concepts are merged. Ontology O_I is following:

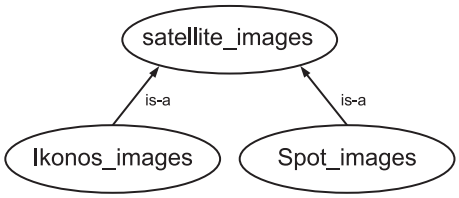


Figure 2: Ontology O_I

Ontology O_I contains three classes: `satellite_images`, `Ikonos_images`, and `Spot_images`. Images from Ikonos and images from Spot are both satellite images, so there is hierarchical class - subclass relationship between `satellite_images` and `Ikonos_images` and between `satellite_images` and `Spot_images`. `Ikonos_images` and `Spot_images` cannot be merged into one class, because it refers to different satellite images. With the knowledge of class properties semantics, there can be seen property - subproperty relationship between a global property and a relevant local property, for example `date` and `date_`. Moreover, if there were the same integrity constraints on each property from the pair, the properties can be merged and connected as equivalent. □

With a data integration system, a user poses his query on the global view in terms of the global view. In order to execute the query over the sources, where data are stored, query processing is needed. There are two approaches to query processing. The first one is query rewriting - a query is decomposed to parts referring to local sources and reformulated to be expressed in local source background. The other one is query answering - it do not pose any limitations on how a query is processed, the only goal is to exploit all possible information to compute the answer, for example find the set of data such that the knowledge logically implies that it is an answer to the query.

With mapping expressed in an ontology, for query rewriting, it is possible to adopt a rule well known in object-oriented world: a child can substitute his parent. If we are looking for all instances of class C that have property $P = x$, the query is

$$q := C(P = x).$$

Using ontology O_I, is-a hierarchy relationships give a means how to rewrite the query with respect to a specific local source. If C is not a concept of the local source schema, class C in the query is replaced with its nearest subclass C' in the is-a hierarchy. This is recursively repeated until a concept is founded in the specific local source schema, or there are no more subclasses - there is no answer. The same rule as for classes can be adopted also for properties, and the relationship property-subproperty can be employed.

In query answering approach to query processing, the is-a hierarchy is also essential. It expresses that an instance of a node is an instance of all nodes within the path from the root node to this node. Based on this rule, it can be determined if information from a local source can be an answer to the global query.

Example 2 Continuing the simple example of satellite images integration, this example shows query processing. The global view provides satellite images. The query: give all available images taken on 1st January 2001, i.e.

$$q := satellite\_images(date =' 01-01-2001'),$$

is processed as follows: `satellite_images` is not in the concept of any local source, the query is rewritten. `satellite_images` class has two child nodes `Ikonos_images` related to the source 1 and `Spot_images` related to the source 2. The reformulated query has two forms:

$$q_1' := Ikonos\_images(date =' 01-01-2001')$$

and

$$q_2' := Spot\_images(date =' 01-01-2001').$$

Because property date is not in the concept of the source 1, the query q_1 is further rewritten using property-subproperty to

$$q_1'' := Ikonos\_images(date\_ =' 01-01-2001').$$

The query q_1'' is executed over the source 1. Analogously, the query q_2' is rewritten and executed over the source 2. □

Compared with two basic approaches of mapping specification in a mediation data integration, an ontology-based approach is similar to LAV integration in a way, that the global schema is specified independently from the sources. Another similarity can be found in extending the system. When a new source is added, the ontology of the integration system O_I is enriched with a new source ontology and further possible relationships to previous version of O_I. A difference between LAV and GAV and the ontology-based integration system is in the case of a change in the layer of local or global source schemas. In case of using ontologies, the ontology of integration system is enriched with the new state. It is not needed to change any earlier part of the ontology, or even to remove some part. No other change is needed.

5. Conclusion

Data integration is a task of combining data from different data sources and enabling a user to process them as one whole. There are two classical ways of designing an integration system providing a global virtual view over the sources: GAV and LAV approaches. Both are based on definition of connection between the global view and the local sources via mappings. However, with a Semantic Web idea, there are also other possibilities that can be used. An integration system described in this paper is based on ontologies of the sources. An ontology of the integration system is defined, and it is consequently used for data query processing.

References

[1] J. D. Ullman, "Information integration using logical views", *Theoretical Computer Science* 239 (2000), pp. 189-210.

[2] M. Lenzerini, "Data Integration: A Theoretical Perspective", In Proceedings of the *21st ACM SIGMOD - SIGACT - SIGART symposium on Principles of database systems*, ACM Press, pp. 233-246, 2002.

[3] A. Cali, D. Calvanese, G. De Giacomo, and M. Lenzerini, "On the Expressive Power of Data Integration Systems", In Proceedings of the *21st Int. Conf. On Conceptual Modeling (ER 2002)*, LNCS 2503, Springer, pp. 338-350, 2002.

[4] M. Friedman, A. Levy, and T. Millstein, "Navigational plans for data integration", In Proceedings of the *16th Nat. Conf. On Artificial Intelligence (AAAI'99)*, AAAI Press, pp. 67-73, 1999.

[5] T. Berners-Lee, J. Hendler and O. Lassila, "The Semantic Web", *Scientific American*, vol. 284, 5, pp. 35-43, 2001.

[6] M.-R. Koivunen and E. Miller, "W3C Semantic Web Activity", in the proceedings of the *Semantic Web Kick/off Seminar*, Finland, 2001.

[7] W3C (WWW Consortium), `http://www.w3.org`.

[8] Extensible Markup Language (XML), `http://www.w3.org/XML/`.

[9] Resource Description Framework (RDF), `http://www.w3.org/RDF/`.

[10] RDF Vocabulary Description Language 1.0: RDF Schema, *W3C Recommendation*, `http://www.w3.org/TR/2004/REC-rdf-schema-20040210`, February, 2004.

[11] Y. Ding, D. Fensel, M. Klein, and B. Omelayenko, "The semantic web: yet another hip?", *Data & Knowledge Engineering* 41 (2002), pp. 205-227.

[12] Web Ontology Language (OWL), `http://www.w3.org/2004/OWL`.

[13] K. Kotis, G. A. Vouros, and K. Stergiou, "Towards automatic merging of domain ontologies: The HCONE-merge approach", *Web Semantics: Science, Services and Agents on the World Wide Web* 4 (2006), pp. 60-79.

[14] H. T. Uitermark, P. J. M. van Oosterom, N. J. I. Mars, and M. Molenaar, "Ontology-based integration of topographic data sets", *International Journal of Applied Earth Observation and Geoinformation* 7 (2005), pp. 97-106.

[15] S. Bergamaschi, S. Castano, M. Vincini, and D. Beneventano, "Semantic integration of heterogeneous information sources", *Data & Knowledge Engineering* 36 (2001), pp. 189-210.

[16] C.-Y. Lee and V.-W. Soo, "The conflict detection and resolution in knowledge merging for image annotation", *Information Processing and Management* 42 (2006), pp. 1030-1055.

Reliability of Educational Tests

Post-Graduate Student:
MGR. PATRÍCIA MARTINKOVÁ
Department of Medical Informatics
Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic
martinkova@euromise.cz

Supervisor:
DOC. RNDR. KAREL ZVÁRA, CSC.
Department of Probability and Mathematical Statistics
Faculty of Mathematics and Physics
Charles University in Prague
Sokolovská 83
186 75 Praha 8
Czech Republic
karel.zvara@mff.cuni.cz

Field of Study:
Probability and Mathematical Statistics

The work was supported by the grant 1M06014 of the Ministry of Education of the Czech Republic.
It contains results published in [1], some recent results were added.

Abstract

The paper deals with reliability of measurements in the context of multiple-item testing instruments, such as educational tests. We concentrate on popular characteristic widely used for estimation of reliability called *Cronbach's alpha*, which is suited for normally distributed error term. Further we discuss modifications of Cronbach's alpha for the case of dichotomous (true-false) scoring.

1. Reliability

When describing the reliability of measurement, it is usually assumed that the measurement Y is composed out of two random variables: an unobservable true value T and an error term e,

$$Y = T + e.$$

The error term is supposed to have a zero mean $E(e) = 0$, a positive variance, and to be independent from the true value T. Therefore:

$$\text{var}\,(Y) = \text{var}\,(T) + \text{var}\,(e).$$

The *reliability* of such measurement is defined by:

$$R = \frac{\text{var}\,(T)}{\text{var}\,(Y)} = 1 - \frac{\text{var}\,(e)}{\text{var}\,(Y)} \tag{1}$$

and it compares variability of the error term with the variability of measured property. The smaller the error variance relative to the observed score variance, the more reliable is the measurement. Thus, the measurement is considered to be reliable when the value of reliability is close to 1.

Here, we should point out that reliability is sample-dependent. Therefore a certain test can have a different reliability when given to a population with a high variability of tested knowledge than when given to a population with a low variability of the knowledge.

The following simple lemmas give us a natural interpretation of the reliability.

Lemma 1.1 *Having two independent measurements $Y_1 = T + e_1, Y_2 = T + e_2$ of the same property T, where var (e_1) = var (e_2), the reliability can be expressed as the correlation between these two measurements, R = corr (Y_1, Y_2).*

Proof:

$$\text{corr}\,(T + e_1, T + e_2) = \frac{\text{cov}\,(T + e_1, T + e_2)}{\sqrt{\text{var}^2(Y)}} = \frac{\text{cov}\,(T, T) + 0}{\text{var}\,(Y)} = \frac{\text{var}\,(T)}{\text{var}\,(Y)} = R. \qquad \square$$

In terms of educational tests, the reliability reflects to what extent it gives the same result when taken repeatedly by the same person under the same conditions.

Lemma 1.2 *The reliability can be expressed as the squared value of the correlation between the observed score and the true score, corr* $^2(Y,T)$.

Proof:

$$\text{corr}^2(Y,T) = \frac{\text{cov}^2(T+e,T)}{\text{var}(Y)\text{var}(T)} = \frac{\text{var}^2(T)}{\text{var}(Y)\text{var}(T)} = \frac{\text{var}(T)}{\text{var}(Y)} = R.$$

□

Thus, the reliability of an educational test measures the strength of the relationship between the score reached by a student and his/her true knowledge.

Unfortunately, none of these representations is useful when estimating the reliability of educational tests because they cannot be directly estimated from the observed data. We cannot estimate the error variance var (e), the true score T, nor the knowledge of a student by the same test twice and independently. Therefore, when estimating the reliability of an educational test, we mostly take into account a fact that such a test is a composite measurement.

2. Reliability of composite measurement

We consider the problem of measuring the reliability of multiple-item testing instrument, such as in educational test. Consider a series of items $Y_j = T_j + e_j$, for $j = 1,\ldots,m$, where the error terms e_j are mutually independent and independent on the true scores T_k for $k = 1,\ldots,m$, having the same variance var $(e_j) = \sigma_e^2$, and mean $\mathrm{E}\, e_j = 0$. The observed overall score of the m items is given by $Y = Y_1, + \cdots + Y_m$ and the unobservable overall true score is given by $T = T_1, + \cdots + T_m$. The reliability of such a composite measurement is defined by (1) and with regard to the above mentioned assumptions can further be expressed as:

$$R_m = \frac{\text{var}(T)}{\text{var}(Y)} = \frac{\text{var}(T)}{\text{var}(T) + \text{var}(\sum e_j)} = \frac{\text{var}(T)}{\text{var}(T) + m\sigma_e^2}. \tag{2}$$

The next lemma gives a relationship between reliability of a composite measurement and reliability of an item in one special case:

Lemma 2.1 *If for the items' true score the following holds simultaneously:*

$$\textit{var}(T_1) = \cdots = \textit{var}(T_m) = \sigma_T^2$$
$$\textit{corr}(T_j, T_k) = 1, \qquad j,k = 1,\ldots,m,$$

then all the reliabilities R_1 *of the items are equal and the reliability of the whole test can be expressed in* ***Spearman-Brown formula****:*

$$R_m = \frac{mR_1}{1+(m-1)R_1} \tag{3}$$

Proof:

$$\begin{aligned}
\text{var}\left(\sum_{j=1}^m T_j\right) &= \sum_{j=1}^m \text{var}(T_j) + \sum\sum_{j\neq k} \text{cov}(T_j, T_k) = \\
&= m\sigma_T^2 + m(m-1)\sigma_T^2 = m^2\sigma_T^2, \\
\text{var}\left(\sum_{j=1}^m Y_j\right) &= \text{var}\left(\sum_{j=1}^m T_j\right) + \text{var}\left(\sum_{j=1}^m e_j\right) = m^2\sigma_T^2 + m\sigma_e^2.
\end{aligned}$$

Therefore

$$\begin{aligned} R_m &= \frac{\text{var}\left(\sum_{j=1}^{m} T_j\right)}{\text{var}\left(\sum_{j=1}^{m} Y_j\right)} = \frac{m^2\sigma_T^2}{m^2\sigma_T^2 + m\sigma_e^2} = \frac{m\frac{\sigma_T^2}{\sigma_T^2+\sigma_e^2}}{1+(m-1)\frac{\sigma_T^2}{\sigma_T^2+\sigma_e^2}} \\ &= \frac{mR_1}{1+(m-1)R_1}. \end{aligned}$$

□

Related to this lemma is a fact, that reliability of an educational test is dependent on the number of its items. Therefore, by adding suitable items to the test, the reliability could approach as close to 1 as we would desire. When comparing reliabilities of two educational tests, which in principle can't have the same number of items, we should bear this property of reliability in mind.

3. Cronbach's Alpha

As a measure of reliability in classical test theory, Cronbach [2] proposed the coefficient alpha. This characteristic estimates the consistency between items in a test and it is defined as:

$$\alpha_{CR} = \frac{m}{m-1}\frac{\text{var}\,(Y) - \sum_j \text{var}\,(Y_j)}{\text{var}\,(Y)} = \frac{m}{m-1}\frac{\sum\sum_{j\neq k}\sigma_{jk}}{\sum\sum_{j,k}\sigma_{jk}}, \tag{4}$$

where σ_{jk} is the covariance of the pair (Y_j, Y_k). Novick and Lewis [3] has shown that Cronbach's alpha is always a lower bound of the reliability

$$\alpha_{CR} \leq R$$

and is equal to reliability only if the conditions of Lemma 2.1 are fulfilled.

A very pleasant property of Cronbach's alpha is the fact that this characteristic is easy to estimate from the data simply by using sample variances and sample covariances instead of their population counterparts in (4). This sample estimate can further be rewritten (for proof see [4]) in terms of the two-way ANOVA as:

$$\hat{\alpha}_{CR} = \frac{MS_T - MS_e}{MS_T} = 1 - \frac{1}{F_T}, \tag{5}$$

where MS_T and MS_E are the mean sums of squares and F_T is statistics widely used for testing the hypothesis $\text{var}\,(T) = 0$ when normality of variables can be assumed.

Notation (5) gives important properties of our estimate:

- $\hat{\alpha}$ can take values between $-\infty$ and 1, although only positive values make sense for reliability.
- The greater the estimate of reliability is, the better the educational test can distinguish between the students. This points out the fact, that Cronbach's alpha was designed as a coefficient of internal consistency.
- The estimate equals one, if and only if there exist constants $a_i, b_j, i = 1, \ldots, n, j = 1, \ldots, m$, so that the score reached by the i-th student in the j-th item can be written as $a_i + b_j$. This means that in this case, to get all the information about students, one item would be enough. Therefore, when getting too high an estimate of Cronbach's alpha, one should actually think of lowering the number of items.

4. Cronbach's alpha for dichotomous items

In fact, Cronbach's alpha was designed as a generalization of the so called Kuder-Richardson formula 20 for dichotomous scoring, already proposed in 1937 in [6]:

$$\hat{\alpha} = \frac{m}{m-1}\frac{s^2 - \sum_{j=1}^{m} p_j(1-p_j)}{s^2}, \tag{6}$$

where p_j is a relative frequency of correct answers to the jth item and s^2 is a sample estimate of the variance of total scores. One can easily see that (6) can be obtained when computing the sample estimate of Cronbach's alpha (4) in the case of dichotomous scoring, where $\hat{\mathrm{E}}\, Y_j = p_j$ is the proportion of correct answers to the jth question and $\hat{\mathrm{var}}\,(Y_j) = p_j(1-p_j)$.

Nevertheless, with dichotomous items, the assumptions of analysis of variance are violated. The scores cannot be assumed to have normal distribution, and moreover, the variance is dependent on the mean value. Therefore it is a matter of question to what extent is this estimate appropriate at all.

4.1. Proposed modifications of Cronbach's alpha estimate

Formula (5) led Zvara [5] to the idea of modifying Cronbach's alpha for the case of binary outcomes by replacing F_T by statistics used for testing the hypothesis $H_0 : \mathrm{var}\,(T) = 0$ in logistic regression. This is equal to testing the submodel B where the score Y_{ij} depends only on the test item (and doesn't depend on the student's ability) against the model A+B where the score Y_{ij} depends on the student and on the test item. Appropriate statistics is the difference of deviance in the submodel and in the model $X^2 = D(B) - D(A+B)$, which has under the null hypothesis the $\chi^2(n-1)$ distribution. Therefore, the proposed estimate is:

$$\hat{\alpha}_{log} = 1 - \frac{n-1}{X^2}. \tag{7}$$

In this work we are trying to justify the estimate (7), so far called the *logistic estimate of Cronbach's alpha* or shortly *logistic alpha*, and to demonstrate its qualities by simulations.

5. Extended beta-binomial model

The model used most often for describing items with dichotomous scoring is the logit-normal model called Rasch model [7], [8]. In this model the probability of a correct response of person i on item j is given by:

$$\mathrm{P}\,(Y_{ij} = y_{ij}; \pi_i, \delta_j) = \frac{\exp[y_{ij}(\pi_i + \delta_j)]}{1 + \exp(\pi_i + \delta_j)}, \tag{8}$$

where π_i describes the level of ability of person i and δ_j is an unknown parameter describing the difficulty of item j.

Evaluating the true reliability of a composite measurement of items which obey Rasch model with certain parameters $\pi_i, i = 1, \ldots, n$, and $\delta_j, j = 1, \ldots, m$ is a difficult task. That is why for simulations we propose the *extended beta-binomial model*, where calculating the true reliability is tractable (see (9)). The motivation of this model is following:

An often used model in reliability studies of binary data (see for example [9], [10]) is the *beta-binomial model*. In this model we assume, that the probability of success π_i varies over subjects $i = 1, \ldots, n$ according to a beta distribution with parameters a and b, and, conditional to this probability, the total score Y_i of the ith person is binomially distributed. Choice of beta distribution for π_i is logical since it is a flexible distribution and leads to mathematically tractable results. A pleasant property of the beta-binomial model is the fact, that the first two moments for the total score are easy to compute:

$$\mathrm{E}\,(Y) = n\mu = n\frac{a}{a+b}$$

$$\mathrm{var}\,(Y) = n\mu(1-\mu)\left[1 + (n-1)\frac{\theta}{1+\theta}\right],$$

where μ is the marginal probability of success for any individual, $\theta = \frac{1}{a+b}$ and $\frac{\theta}{1+\theta} = \rho$ is the intraclass correlation $\mathrm{corr}\,(Y_{ij}, Y_{ik})(j \neq l)$ common for any subject and any pair of responses.

An unpleasant property of this model for our situation is the fact that it does not allow for different difficulties of items. Hand in hand with this goes the common-correlation structure which is impossible in our case.

When trying to extend for different difficulties of items and yet preserve the structure of the beta-binomial model, we can think of the following model: We assume again, that the probability of success π_i varies over subjects $i = 1, \ldots, n$ according to a beta distribution with parameters a and b. We qualify the impact of the difficulty of the jth item by a small number δ_j, assuming that $\sum_{j=1}^{m} \delta_j = 0$. When parameters a, b are large enough, there is a slight danger that the sums $\pi_i + \delta_j$ get outside the interval $(0, 1)$. Therefore, $Y_{i1}, \ldots, Y_{im}$ are for a given π_i independent random variables with alternative distribution alt$(\pi_i + \delta_j)$. The total scores Y_i are sums of such random variables.

5.1. Properties of Y_{ij} in the extended beta-binomial model

For conditional mean and variance, it holds

$$\begin{aligned}
\mathrm{E}\,(Y_{ij}|\pi_i) &= \mathrm{E}\,(Y_{ij}^2|\pi_i) = \mathrm{P}\,(Y_{ij} = 1|\pi_i) = \pi_i + \delta_j, \\
\mathrm{var}\,(Y_{ij}|\pi_i) &= \mathrm{E}\,(Y_{ij}^2|\pi_i) - (\mathrm{E}\,(Y_{ij}|\pi_i))^2 = (\pi_i + \delta_j)(1 - (\pi_i + \delta_j)).
\end{aligned}$$

Therefore the unconditional mean is

$$\mathrm{E}\,(Y_{ij}) = \mathrm{E}\,\mathrm{E}\,(Y_{ij}|\pi_i) = \frac{a}{a+b} + \delta_j = \mu + \delta_j,$$

where we assigned $\mu = a/(a+b)$ for the mean value of the beta distribution. For the unconditional variance it holds:

$$\begin{aligned}
\mathrm{var}\,(Y_{ij}) &= \mathrm{var}\,(\mathrm{E}\,(Y_{ij}|\pi_i)) + \mathrm{E}\,(\mathrm{var}\,(Y_{ij}|\pi_i)) \\
&= \mathrm{var}\,(\pi_i + \delta_j) + \mathrm{E}\,((\pi_i + \delta_j)(1 - (\pi_i + \delta_j))) \\
&= \frac{ab}{(a+b)^2(a+b+1)} + \int_0^1 (\pi + \delta_j)(1 - (\pi + \delta_j)) \frac{1}{\mathrm{B}(a,b)} \pi^{a-1}(1-\pi)^{b-1} \mathrm{d}\pi \\
&= \frac{ab}{(a+b)^2(a+b+1)} + (\mu + \delta_j) - \frac{a(a+1)}{(a+b)(a+b+1)} - 2\delta_j \frac{a}{a+b} - \delta_j^2 \\
&= \mu(1-\mu) + \delta_j(1 - 2\mu - \delta_j).
\end{aligned}$$

Because $\rho = \frac{\theta}{1+\theta} = \frac{1}{a+b+1}$, the covariance of variables Y_{ij}, Y_{it} for $j \neq t$ equals

$$\begin{aligned}
\mathrm{cov}\,(Y_{ij}, Y_{it}) &= \mathrm{cov}\,(\mathrm{E}\,(Y_{ij}|\pi_i), \mathrm{E}\,(Y_{it}|\pi_i)) \\
&= \mathrm{cov}\,(\pi_i + \delta_j, \pi_i + \delta_t) = \mathrm{var}\,(\pi_i) \\
&= \frac{ab}{(a+b)^2(a+b+1)} = \rho\mu(1-\mu),
\end{aligned}$$

Let's define

$$C_j = 1 + \delta_j \frac{1 - 2\mu - \delta_j}{\mu(1-\mu)}.$$

Then the correlation between Y_{ij} and Y_{it} for $j \neq t$ is

$$\mathrm{corr}\,(Y_{ij}, Y_{it}) = \frac{\mathrm{cov}\,(Y_{ij}, Y_{it})}{\sqrt{\mathrm{var}\,Y_{ij}\,\mathrm{var}\,Y_{it}}} = \rho \frac{1}{\sqrt{C_j C_t}}.$$

For constant difficulties of items $\delta_j = 0$ we get the common correlation structure, corr $(Y_{ij}, Y_{it}) = \rho$. For unequal difficulties of items it is natural to assume $a = b$ (to assume symmetric distribution of knowledge), therefore $\mu = 1/2$. In this case $C_j = 1 - 4\delta_j^2$, thus the impact of $\delta < 1$ is small.

5.2. Properties of total scores Y_i in the extended beta-binomial model

The total score of the ith student is the total number of correctly answered items $Y_i = \sum_{j=1}^{m} Y_{ij}$. We get

$$\begin{aligned}
\mathrm{E}\,(Y_i|\pi_i) &= m\frac{1}{m}\sum_{j=1}^{m}(\pi_i+\delta_j) = m\pi_i,\\
\mathrm{var}\,(Y_i|\pi_i) &= \mathrm{var}\,(\sum_{j=1}^{m} Y_{ij}|\pi_i) = \sum_{j=1}^{m}(\pi_i+\delta_j)(1-(\pi_i+\delta_j) = m\frac{1}{m}\sum_{j=1}^{m}(\pi_i+\delta_j) - m\frac{1}{m}\sum_{j=1}^{m}(\pi_i+\delta_j)^2\\
&= m\left(\pi_i - \left(\frac{1}{m}\sum_{j=1}^{m}(\pi_i+\delta_j)\right)^2\right) - m\left(\frac{1}{m}\sum_{j=1}^{m}(\pi_i+\delta_j)^2 - \left(\frac{1}{m}\sum_{j=1}^{m}(\pi_i+\delta_j)\right)^2\right)\\
&= m\pi_i(1-\pi_i) - m\kappa_\delta,
\end{aligned}$$

where $\kappa_\delta = \frac{1}{m}\sum_{j=1}^{m}\delta_j^2$. Therefore it holds:

$$\begin{aligned}
\mathrm{E}\,(Y_i) &= \mathrm{E}\,\mathrm{E}\,(Y_i|\pi_i) = m\mathrm{E}\,(\pi_i) = m\frac{a}{a+b} = m\mu,\\
\mathrm{var}\,(Y_i) &= \mathrm{var}\,(m\pi_i) + \mathrm{E}\,(m\pi_i(1-\pi_i) - m\kappa_\delta) = m^2\mathrm{var}\,(\pi_i) + m\mathrm{E}\,(\pi_i) - m\mathrm{E}\,(\pi_i^2) - m\kappa_\delta)\\
&= m^2\frac{ab}{(a+b)^2(a+b+1)} + m\frac{a}{a+b} - m\frac{a(a+1)}{(a+b)(a+b+1)} - m\kappa_\delta\\
&= m\mu(1-\mu)(1+(m-1)\rho) - m\kappa_\delta.
\end{aligned}$$

Finally, we are getting to the **reliability of the total score** Y_j in the extended binomial model. We define it according to [11] as a fraction of variability between students (variability of conditional mean values $\mathrm{E}\,(Y_i|\pi_i)$) and variability of students' total scores Y_i:

$$\begin{aligned}
R_m &= \frac{\mathrm{var}\,(\mathrm{E}\,(Y_i|\pi_i))}{\mathrm{var}\,(Y_i)} = \frac{\mathrm{var}\,(m\pi_i)}{\mathrm{var}\,(Y_i)} = \frac{m^2\mu(1-\mu)\rho}{m\mu(1-\mu)(1+(m-1)\rho) - m\kappa_\delta}\\
&= \frac{m\rho}{1+(m-1)\rho - \frac{\kappa_\delta}{\mu(1-\mu)}}.
\end{aligned} \tag{9}$$

When the difficulties of items are all equal $\delta_j = 0$, we get the well known Spearman-Brown formula (3). For unequal difficulties of items, the reliability of total scores is a bit larger.

Formula (9) is very important for simulations. For the given parameters of beta-binomial distribution a, b, and for the given difficulties δ_j, $j = 1, \ldots, m$ we can calculate the true reliability R_m and compare it with estimates calculated from simulated data.

6. Simulations

So far, a single simulation was done. We investigated the behavior of the classical and logistic Cronbach's alpha estimator in the extended beta-binomial distribution via simulation for number of items $k = 11$, number of students $n = 20$ and items' difficulties δ_j equidistantly distributed between -0.1 and 0.1. The parameters $a = b$ of the beta-binomial distribution were chosen from the interval $\langle 1, 15\rangle$ with step 0.2. For each simulation we generated 100 data sets and computed bias and mean squared error of the classical estimates of Cronbach's alpha and of the logistic estimates of Cronbach's alpha. For each out of 71 possible values of parameters $a = b$, also the theoretical value of reliability was evaluated, using the equation (9).

In Figure 1, the bias and mean squared error of classical and logistic estimate of Cronbach's alpha is shown for different values of the true reliability.

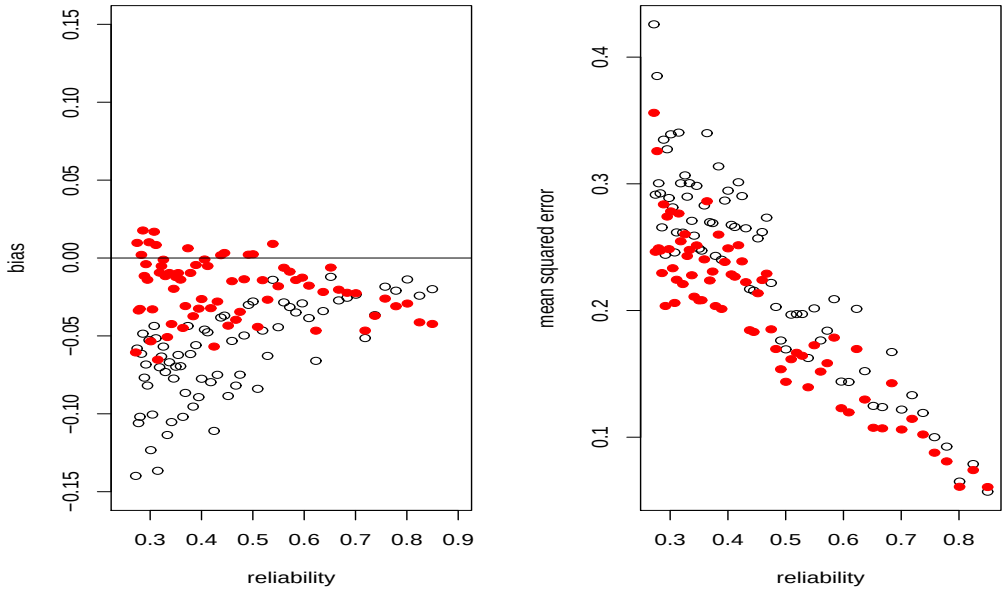


Figure 1: The bias and mean squared error of classical (circle) and proposed logistic (solid circle) estimate of α_{CR}

According to Figure 1, the proposed logistic alpha performs as an estimate of reliability better than the classical Cronbach's alpha estimate in the extended beta-binomial model. The logistic alpha tends to give worse results only for high true reliabilities, thus for small a, b, which is the case of high probability of cutting in the extended beta-binomial model.

7. Conclusions and Discussion

According to our simulations, the proposed logistic estimate of Cronbach's alpha performs better for binary data of the extended beta-binomial model than the classical Cronbach's alpha estimate.

When going through section 5, one can conclude, that there is no need for beta distribution in the extended model to get the same formula for true reliability of total scores Y_i. Therefore, more complex simulations in this class of models should be done.

Also, the remaining task is to justify the proposed class of models for the real data (would Hosmer-Lemeshow goodness-of-fit test work?). Or even better to justify the proposed estimate (7) for the Rasch model (8).

For testing the hypothesis $H_0 : \mathrm{var}\,(T) = 0$, there also exist other statistics besides difference of deviances. Therefore, other modifications of classical Cronbach's alpha estimate could also be defined and compared with logistic alpha discussed in this article.

References

[1] P. Martinková, K. Zvára jr., J. Zvárová, K. Zvára "The New Features of the ExaMe Evaluation System and Reliability of Its Fixed Tests", *Methods of Information in Medicine*, vol. 45, pp. 310–315, 2006.

[2] L. J. Cronbach "Coefficient Alpha and the internal structure of tests", *Psychometrika*, vol. 16, pp. 297–334, 1951.

[3] M. R. Novick, Ch. Lewis *Coefficient Alpha and the Reliability of Composite Measurements*, Educational Testing Service, Princeton, New Jersey, 1966.

[4] P. Rexová, *Spolehlivost měření* [Reliability of measurements, In Czech] Diploma thesis. Department of Probability and Mathematical Statistics, Charles University, Prague 2003.

[5] K. Zvára "Měření reliability aneb bacha na Cronbacha". *Statistický bulletin*, vol. 13, pp. 13–20, 2002.

[6] G. Kuder, M. Richardson "The theory of estimation of test reliability". *Psychometrika*, vol. 2, pp. 151-160, 1937.

[7] Rasch, G., *Probabilistic Models for Some Intelligence and Attainment Tests*, The Danish Institute of Educational Research, 1960.

[8] P. Rexová "Item Analysis of Educational Tests in System ExaMe". Doktorandský den '04. ICS AS CR, 2004.

[9] M. S. Ridout, C. G. B. Demeterio, D. Firth "Estimating intraclass correlation for binary data". *Biometrics*, vol. 55, pp. 137–148, 1999.

[10] G. Zou, A. Donner "Confidence Interval Estimation of the Intraclass Correlation Coefficient for Binary Outcome Data". *Biometrics*, vol. 60, pp. 807–811, 2004.

[11] D. Commenges, H. Jacqmin "The Intraclass Correlation Coefficient Distribution-Free Definition and Test". *Biometrics*, vol. 50, pp. 517–526, 1994.

Data Models' Synchronization in MUDRlite EHR

Post-Graduate Student:
MGR. MIROSLAV NAGY

Department of Medical Informatics
Instutite of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8

Czech Republic

nagy@euromise.cz

Supervisor:
RNDR. ANTONÍN ŘÍHA, CSC.

Department of Medical Informatics
Instutite of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8

Czech Republic

riha@euromise.cz

Field of Study:
Biomedical Informatics

The work was supported by the project number 1ET200300413 of the Academy of Sciences of the Czech Republic.

Abstract

Contemporary electronic health record (EHR) applications incorporate comprehensive medical knowledge that is being updated very fast. Even EHRs that are built considering best practices vary slightly in their data models (DM) over time. There are various software tools supporting "DM synchronizations" on the market. Within this paper we describe special requirements that are to be set on these tools to be successfully applicable for synchronization of DMs in healthcare domain. As the available tools do not meet the requirements well we also present a new open source DM synchronization tool Schemagic that is particularly suitable for these synchronizations. Schemagic was developed as a supportive tool within the project Information Technologies for Shared Health Care. Several partners from academy as well as from industry joined for collaboration on this project in order to cooperate on new approaches to support sharing of medical data and knowledge among heterogeneous information systems. The Schemagic synchronization tool has been successfully tested while formalizing the clinical contents of EHRs in dental medicine. A rapid evolution of the underlying DM that started with a few clinical concepts and evolved stepwise into hundreds of attributes and relations has confirmed that Schemagic considerably multiplies the power and scalability of EHR systems and significantly simplifies further research and development.

1. Introduction

Variation in the use of clinical resources, outcomes, costs, access to health care, clinical content of electronic health records (EHR), and quality of provided health care are well recognized, ever present aspects. It is a phenomenon that affects all sectors of the health care delivery process that is important to clinicians, administrators, and patients [1]. New generation EHRs should be based on an architecture flexible enough to last several generations and allowing for improvements, but without starting from scratch [2].

An EHR for integrated care is defined as a repository of information regarding the health status of a subject of care in computer processable form, stored and transmitted securely, and accessible by multiple authorized users [3]. It has a standardized or commonly agreed logical information model which is independent of EHR systems. Its primary purpose is the support of continuing, efficient and quality integrated health care and it contains information which is retrospective, concurrent and prospective.

The clinical content (i.e., collected concepts) of an EHR for integrated care evolves over time. There remains a wide void between the basic definition and the ideal existence of a standardized EHR. The reality is that we lack a common data model (DM), a common set of data elements, common vocabulary, and a common set of scenarios [4]; however, there is some hope in the CEN prEN 13606 becoming an EU standard in the near future. Still, even the EHRs built considering best practices vary slightly in their

DMs over time, which is an apparent consequence of the variation mentioned at the beginning of this paper. Moreover, contemporary EHR systems incorporate comprehensive medical knowledge that is being updated regularly and that also might force the DMs to slightly evolve. A seamless support of information flow for increasingly distributed healthcare processes requires integrating of heterogeneous IT systems into a comprehensive distributed information system. Different standards contribute to ease this integration [5]; however, a kind of synchronization of the underlying DMs is often enforced.

Within the project Information Technologies for Shared Health Care several partners from academy as well as from industry joined in order to cooperate on new approaches to the EHR design with the main goal of solving issues connected with sharing medical data among heterogeneous hospital information systems (HIS) and EHR applications. Together, they are preparing solutions to model various EHRs and HISs using worldwide-used standards, based mainly on the HL7 version 3 specification [6]. The MUDRLite EHR system [7] is among pilot applications chosen to test new approaches. It provides interfaces to include user-defined modules and components. These interfaces enable to develop and integrate special components to share data among other systems based on a defined communication standard. MUDRLite architecture is based on two tiers. The first one is a relational database and the second one is a MUDRLite User Interface. The database schema corresponds to particular needs and varies therefore in different medical environments. An important part of the project was to incorporate and structuralize clinical information in dental medicine using the Dental Cross component [8], [9].

Evolution of the incorporated dental medicine knowledge obviously forced the evolution of the underlying DM over time. This paper discusses issues we faced and requirements we analyzed while synchronizing DMs in order to support continuous health care. Approaches we came up with can not only be applied to EHRs in dental medicine but to information systems in healthcare in general terms.

2. Synchronization of Data Models in Healthcare-oriented Information Systems

The rapid progress in database systems research over the past couple of decades has resulted in the evolution of diverse database environments. Consequently, developing a declarative approach to schema integration in the context of heterogeneous database systems was major goal of various research activities [10]. Lately, relational databases with object-oriented extensions became a standard in the data management field. As Entity-Relationship (ER) modelling approach became a widely used methodology of designing database models and schemas [11] various issues connected with the synchronization of DMs and maintenance of data integrity developed to an extensively investigated issue [12].

Resulting from the ongoing research and answering to the presented demand various software tools supporting DM comparisons and synchronizations have been produced, e.g., [13], [14], [15], or [16] to name but a few. All the tools are generally intuitive to visually compare and synchronize two databases. They scan both databases and provide the combined schema tree view, with all differences marked. However, even that these tools are fast and easy to operate we have realized that they do not meet all requirements that are to be implemented while synchronizing of DMs in information systems in medicine. Basic functionality of a DM synchronizing tool consists of being capable to read metadata of relational database schemas, comparing two database schemas in order to find changes that are to be solved during the synchronization process, and generation of the synchronizing SQL script. However, further analysis shows that a successful tool synchronizing DMs in healthcare domain should not only provide the basic functionality but also it needs to fulfil other requirements stated as follows:

- *Universality.* As information systems in healthcare domain are increasingly distributed over various heterogeneous IT systems the tool should be universal enough to be run on various operating system platforms (e.g., MS Windows, various Linux based systems, Mac OS) and capable of co-operating with many database system, e.g., Oracle, MS SQL Server, DB2, PostgreSQL, MySQL.
- *No data loss.* The tool must not affect (delete or change) any data stored in the database that is being synchronized. This is especially important in the healthcare domain where patients' data are very

sensitive and the importance of this requirement is well recognized. Even though this requirement seems to be obvious, it is often not the case. A simple data loss may be caused by rounding errors while converting between incompatible data types or by truncating large text fields. Less obvious (and thus more dangerous) errors may be introduced while losing of some functional relations and dependencies that exist in the original database and cannot be held in the new model. Best practices of database normalization and DMs design shall always be followed [17].

- *Extended security*. Patient identification is one of the highest and most controversial priorities for the implementation of HISs [18]. There is a broadly shared goal to better understand the long-term health status of patients when addressing their immediate needs, to study the effectiveness of different patterns of care, to investigate the long-term outcomes of proposed interventions through clinical research studies, and to optimize the system of healthcare delivery; however, a misuse of these sensitive information shall always be avoided. Within a HIS it may be often the case that the clinical data are separated from the administrative ones so that database administrators cannot get the full-context information they do not need. A successful tool used to synchronize DMs in healthcare information systems shall support these scenarios.

- *Extensibility*. The tool should be extensible in different ways in order allow the user (developer) adding of new functionality as this may be of significant importance in the continuously evolving medical domain, e.g., the user should be able to extend the set of supported database objects. Synchronizing capabilities ought to be improved or extended in a convenient way, which means, for example, without the necessity of developing pieces of program code (e.g., new Java classes). However, it also should be possible to add new features by adding plug- ins, which is lately a common way of extending contemporary software.

- *Off-line usage*. The schema synchronization capability should be available also in an off-line mode without a direct connection to a particular database at the time of synchronization. This feature is generally required in 24x7x365 systems providing year-round performance with critical consequences when the system goes down. Certainly, HISs are the case. The offline usage supports synchronizing the developers' and customers' schemas before the application deployment.

- *Automation support*. As the synchronization process in distributed environments, e.g., complex healthcare information systems, is usually not a simple one, a command line interface becomes useful. It helps to manipulate schemas of various servers and to synchronize their models by previously well tested scripts.

- *Standards-based design*. In order to ease the end-user usage as well as potential extensibility by third parties the application of standards is always essential.

- *Extensive reporting*. A successful synchronizing tool should be capable to generate a well structured report about the synchronization performed.

- *DM documentation*. The ability of keeping updated documentation about the current DM is essential. This usually does not take much additional effort as the tool needs to learn the DM details to perform the synchronization anyway.

- *National languages support*. Correct handling of database objects named or in any other way linked with terms using language specific characters is an inherent feature that can be essential in localized information systems. Even though this seems obvious it is often not the case.

As mentioned above there are various tools supporting DMs synchronization. However, these tools can mainly serve as good inspirations but they cannot be directly used to synchronize DMs of EHRs and HISs as they do not meet all the requirements well. This was mainly the reason why we have developed a new synchronizing tool in frame of our research aimed to support continuous shared health care.

3. Schemagic – Data Models Synchronization Tool

Schemagic is a tool developed for DM synchronization with the focus on healthcare environment. It was developed in frame of the applied research within the project *Information Technologies for Shared Health Care* number 1ET200300413 of the Academy of Sciences of the Czech Republic. Further description of Schemagic's functionally, features, and usability can be found in [19]. It is based on current state of the art in this field and on ideas gained while studying various synchronization tools with the main goal to meet all the requirements stated especially for DMs synchronization in healthcare domain. The tool is open source and can be used by anyone for free. Its source codes are based on design patterns [20] in order to be easily modifiable and extendable by independent developers in the future.

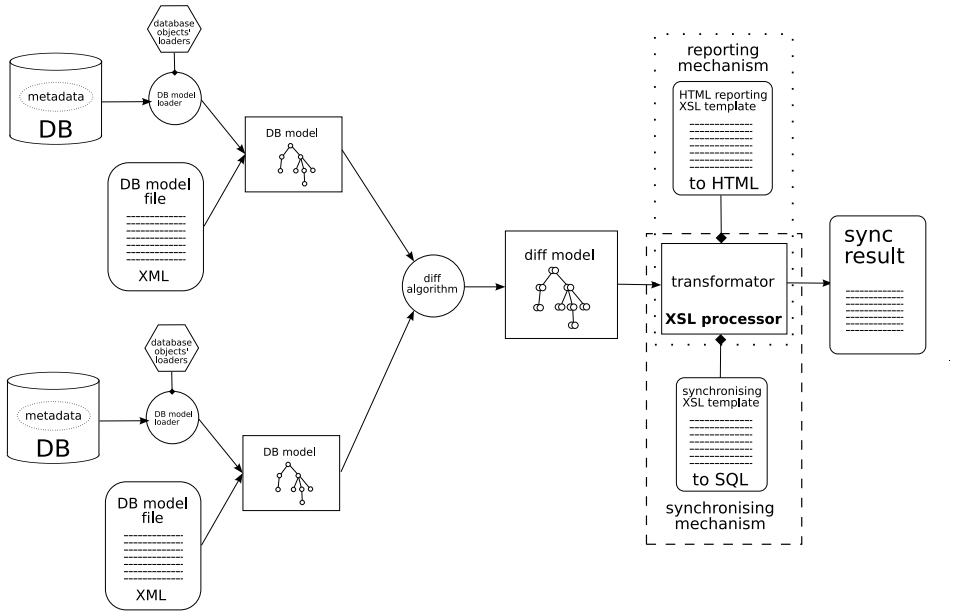


Figure 1: Schemagic's functionality model

The synchronization process in a simplified form is described in Figure 1. During the synchronization source schemas' meta-data are obtained either from a database or from an XML file describing the DM. The next step of the synchronization process is the execution of a specially designed *diff* algorithm that compares both the target and the source schemas and finds differences. The diff algorithm produces a *diff model* that holds all divergences. Final step of the process consists of creation of the synchronization SQL script via XSL transformation [21] mechanism. The script can be run anytime to perform the synchronization itself. The Schemagic tool ready for download as well as its documentation can be found at WWW pages [22].

4. Results

The Schemagic synchronization tool has been successfully tested while applying the MUDRLite EHR into the dental medicine domain [8], [9]. The fact that the detailed structuralization of clinical content of EHRs in dental medicine was a significant part of our research project caused a rapid evolution of the underlying DM that started with a few clinical concepts and evolved stepwise into hundreds of attributes and relations. The Czech national patent application number PV 2005-229 contains some results of this particular research.

Owing mostly to universality of the tool there were no problems with compatibilities of various DMs and their distributed parts that were evolving during the formalization of the clinical EHR content in dental medicine. The capability to generate the DM documentation has made our research and development very convenient. Moreover, we have successfully applied Schemagic while formalizing the clinical content of EHR systems in neurosurgery.

5. Conclusion

There are various tools supporting DMs synchronization available; however, these tools are not particularly suitable for synchronizing of DMs in healthcare domain. We propose a new open source software tool Schemagic that may be used for these synchronizations. Owing to its extensibility there is no problem with support of additional representations of EHR DMs. This tool has been successfully tested while synchronizing various DMs developed within the project Information Technologies for Shared Health Care. Using Schemagic synchronizing tool multiplies the power and scalability of EHR systems and significantly simplifies further research and development. Moreover, the open source based character and the openness of the tool enables additional extensions and modifications of the tool's features according to particular users' needs.

References

[1] T.J. Hannan, "Variation in Health Care – the Roles of the Electronic Medical Record", *Int J Med Inform*, 54 (2), pp. 127–36, 1999.

[2] H. Linden, J. Talmon, H. Tange, G. Boers, A. Hasman, "An architecture for a virtual electronic health record", *Studies in Health Technology and Informatics*, 90: pp. 220–5, 2002.

[3] International Organization for Standardization: ISO/TR 20514:2005, "Health informatics – Electronic health record – Definition, scope and context"

[4] G.C. Liu, J.G. Cooper, K.M. Schoefflerb, W.E. Hammond, "Standards for the Electronic Health Record Emerging from Health Care's Tower of Babel", *Proc AMIA Symp*, pp. 388–92, 2001.

[5] R. Lenz, M. Beyer, K.A. Kuhn, "Semantic Integration in Healthcare Networks", *In: Connecting Medical Informatics and Bio-Informatics: Proceedings of MIE2005 – The XIXth International Congress of the European Federation for Medical Informatics*, vol. 116, pp. 385–90, 2005.

[6] Health Level Seven, http://www.hl7.org (last accessed July 9, 2006).

[7] J. Spidlen, P. Hanzlicek, J. Zvarova, "MUDRLite – Health Record Tailored to Your Particular Needs", *Transformation of Healthcare with Information Technologies*, pp. 202–209, 2004.

[8] J. Spidlen, M. Pies, Z. Teuberova, M. Nagy, P. Hanzlicek, J. Zvarova, T. Dostalova, "MUDRLite – an Electronic Health Record Applied to Dentistry by the Usage of a Dental-Cross Component", *EMBEC'05 Conference*, Prague, 2005.

[9] J. Spidlen, M. Pies, Z. Teuberova, M. Nagy, P. Hanzlicek, J. Zvarova, T. Dostalova, "MUDRLite Components Usage for Sharing EHR Data in Dental Medicine", *In: Sharing Health Information*, Merkantilie, Prague, pp. 83–87, 2005.

[10] L. Lakshmanan, F. Sadri, I. Subramanian, "On the Logical Foundations of Schema Integration and Evolution in Heterogeneous Database Systems", *Proceedings of 2nd Int. Conf. on Deductive and Object-Oriented Databases*, Phoenix, pp. 81–100, 1993.

[11] S. Spaccapietra, S.T. March, Y. Kambayashi, editors, "Conceptual Modeling – ER 2002", *21st International Conference on Conceptual Modeling Tampere*, Heidelberg: Springer, 2002.

[12] C. Batini, M. Lenzerini, S.B. Navathe, "A comparative analysis of methodologies for database schema integration", *ACM Computing Surveys (CSUR)*, vol. 18 n.4, pp. 323–364, 1986.

[13] AdeptSQLTools, "The Adept SQL Diff tool", http://www.adeptsql.com (last accessed Jan 27, 2006).

[14] Spectral Core, "Spectral Core Sync Database", http://www.spectralcore.com (last accessed Jan 27, 2006).

[15] Best SoftTool, "SQLDBCompare", http://bestsofttool.com (last accessed Jan 27, 2006).

[16] Red Gate Software, "SQL Compare", http://www.red-gate.com (last accessed Jan 27, 2006).

[17] P. Rob, C. Coronel, "Database systems, design implementation and management", *Wadsworth Publishing Company*, Belmont, California, 1993.

[18] Board of Directors of the American Medical Informatics Association, "Standards for medical identifiers, codes, and messages needed to create an efficient computer-stored medical record", *The Journal of the American Medical Informatics Association*, 1, pp. 1–7, 1994.

[19] M. Nagy, "Synchronisation of Relational Schemas", *Master Thesis at Faculty of Mathematics and Physics*, Charles University in Prague, 2005.

[20] E. Freeman, "Head First Design Patterns", O'Reilly, 2004.

[21] D. Tidwell, "XSLT", O'Reilly, 2001.

[22] EuroMISE Centre, "Schemagic project", http://schemagic.sourceforge.net (last accessed May 20, 2006).

General Relational Data Model with Preferences

Post-Graduate Student:

RADIM NEDBAL

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic ,

Department of mathematics
Faculty of nuclear science and physical engineering
ČVUT
Trojanova 13
120 00 Praha 2
Czech Republic

radned@seznam.cz

Supervisor:

ING. JÚLIUS ŠTULLER, CSC.

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic

stuller@cs.cas.cz

Field of Study:
Mathematical Engineering

This work was supported by the project 1ET100300419 of the Program Information Society (of the Thematic Program II of the National Research Program of the Czech Republic) "Intelligent Models, Algorithms, Methods and Tools for the Semantic Web Realization", and by the Institutional Research Plan AV0Z10300504 "Computer Science for the Information Society: Models, Algorithms, Applications".

Abstract

The aim of the paper is to present a novel, general approach to preference modelling in the framework of the relational data model. To allow nonmonotonic operations, the preferences are defined between sets of relational instances. The aim is the generalization of the relational algebra that is as minimal as possible, in the sense that the formal fundamentlas of the relational data model are preserved. At the same time, the extended model should be formal enough to provide a sound basis for the investigation of other new preference constructors and operations and for new possible applications.

1. Related Work

Lacroix and Lavency [1] originated the study of preference queries. They proposed an extension of the relational calculus in which preferences for tuples satisfying given logical conditions can be expressed. For instance, one could say: Pick up the tuples of R satisfying $Q \wedge P1 \wedge P2$; if the result is empty, pick the tuples satisfying $Q \wedge P1 \wedge \neg P2$; if the result is empty, pick the tuples satisfying $Q \wedge \neg P1 \wedge P2$. The composition or iteration of preferences, however, is not considered. Neither is addressed the issue of algebraic optimization of preference queries.

Kießling et al. [2, 3, 4, 5, 6, 7] and Chomicki et al. [8, 9, 10, 11] proposed independently a similar framework based on a formal language for formulating preference relations. The embedding (called *Best Match Only* – BMO and *WinNow* – WN respectively) into relational query languages they use is identical. Many possible rewritings for preference queries are presented.

Kießling et al. [2, 6] introduce a number of base preference constructors and their combinators (Pareto and lexicographic composition, intersection, disjoint union, and others). The possibility of having arbitrary constraints in preference formulas is not considered.

The framework of Chomicki et al. [8] emphasizes the view of preferences as strict partial orders and defines preferences more generally as arbitrary logical formulas. Intrinsic and extrinsic classes of preference formulas are studied.

Börzsönyi et al. [12] introduced the skyline operator and described several evaluation methods for this operator. Skyline is a special case of WN and BMO. It is restricted to use an intrinsic preference formula which is a conjunction of pairwise comparisons of corresponding tuple components. Some examples of possible rewritings for skyline queries are given but no general rewriting rules are formulated.

Argawal and Wimmers [13] use quantitative preferences (scoring functions) in queries and focus on the issues arising in combining such preferences. Hristidis et al. [14] explore in this context the problems of efficient query processing using materialized views. As pointed out repeatedly in their paper, the approach based on scoring functions is inherently less expressive than the one based on preference relations. In particular, skyline queries cannot be captured using scoring functions. Moreover, since the quantitative approach is based on comparing the scores of individual tuples under the given scoring functions, the preferences represented in this way have to be intrinsic. In addition, it is not clear how to compose scoring functions to achieve an effect similar to various preference relation composition operations.

A more general approach is proposed in [15], where the relational data model is extended to incorporate partial orderings into data domains. Within the extended model, the partially ordered relational algebra (the PORA) is defined by allowing the ordering predicate to be used in formulae of the selection operation. The PORA expresses exactly the set of all possible relations that are invariant under order-preserving automorphism of databases. This result characterizes the expressiveness of the PORA and justifies the development of Ordered SQL (OSQL) as a query language for ordered databases. OSQL provides users with the capability of capturing the semantics of ordered data in many advanced applications, such as those having temporal or incomplete information.

A similar approach to preference modelling is presented in [16]. A declarative query interface for Web repositories that supports complex expressive Web queries is defined. Such queries have two key characteristics: (i) They view a Web repository simultaneously as a collection of text documents, as a navigable directed graph, and as a set of relational tables storing properties of Web pages (length, URL, title, etc.). (ii) The queries employ application-specific ranking and ordering relationships over pages and links to filter out and retrieve only the "best" query results. The Web repository is modelled in terms of "Web relations". A description of an algebra for expressing complex Web queries is given. The algebra extends traditional relational operations as well as graph navigation operations to uniformly handle plain, ranked, and ordered Web relations. In addition, an overview of the cost-based optimizer and execution engine is presented.

In [17], actual values of an arbitrary attribute are allowed to be partially ordered as well. Accordingly, relational algebra operations, aggregation functions and arithmetic are redefined. Thus, on one side, the expressive power of the classical relational model is preserved, and, at the same time, as the new operations operate on and return ordered relations, information of preference, which is represented by a partial ordering, can be handled. Nevertheless, the redefinition of the relational operations causes loss of some of their common properties. For instance, $A \cap B = A - (A - B)$ does not hold. To rectify this weak point, more general approach is needed.

A comprehensive work on partial order in databases is [18]. It presents the partially ordered sets as the basic construct for modelling data. Collection of algebraic operations for manipulating ordered sets is investigated, and their implementation based on the use of realizers as a data structure is presented. An algorithm for generating realizers for arbitrary finite partial orders is provided.

Various kinds of ordering on powerdomains have been considered in context of modelling incomplete information. A very extensive and general study is provided in [19].

In the context of financial and statistical applications, systems such as SEQUIN [20], SRQL [21], and more recently Aquery [22, 23] have proposed SQL extensions to incorporate ordering.

2. Proposed Approach

All the above approaches miss out defining nonmonotonic operations, e.g. set difference, on relations with partial ordering incorporated into attribute or relation domains. An exception is [17] whose generalized difference operation, however, does not preserve common algebraical equalities. Obviously, a more general model is needed. The proposed solution is to take all relational instances into account instead of mere tuples. To see why, refer to the following paragraph and the following example.

The framework of relational instances with ordered tuples can be understood as a generalization of *fuzzy sets-based model* [24, 25] of uncertainty. The same way as a fuzzy relational instance can be described by its alfa cuts, an instance R^* of an arbitrary relation R with an ordering $\preceq^{R^*}$ can be described by its subset containing appropriate tuples. The intuition suggests considering tuples according to their preferences. That is to say, we take into account the more preferred tuples ahead of those with lower preference. Thus, for each tuple, t_i, we take into account a set, $t_i\!\uparrow$, containing this tuple and all the tuples with higher preferences:

$$t_i\!\uparrow = \{t \mid t \in R^* \ \wedge \ t_i \preceq^{R^*} t\}$$

Note that $t_i\!\uparrow \in \mathscr{I}(R)$. Next, considering a unary relational operation

$$\mathcal{O}_u : \mathscr{I}(R) \to \mathscr{I}(Q),$$

which is a mapping from a set of relational instances of R into a set of relational instances of resulting relation Q, it is applied to each set $t_i\!\uparrow$. Note that $\mathcal{O}_u(t_i\!\uparrow) \in \mathscr{I}(Q)$. Finally, the order $\preceq^{\mathscr{I}(Q)}$ on the resulting collection

$$\{\mathcal{O}_u(t_i\!\uparrow) \mid t_i \in R^*\} \subseteq \mathscr{I}(Q)$$

of sets is to be determined.

Example 1 Consider a set $\{\mathcal{O}_u(t_i\!\uparrow) \mid t_i \in R^*\} \subseteq \mathscr{I}(Q)$ with a relation $\sqsubseteq$ implied by preference $\preceq^{R^*}$ on R^*:

$$\mathcal{O}_u(t_i\!\uparrow) \sqsubseteq \mathcal{O}_u(t_j\!\uparrow) \Leftrightarrow t_i \preceq^{R^*} t_j$$

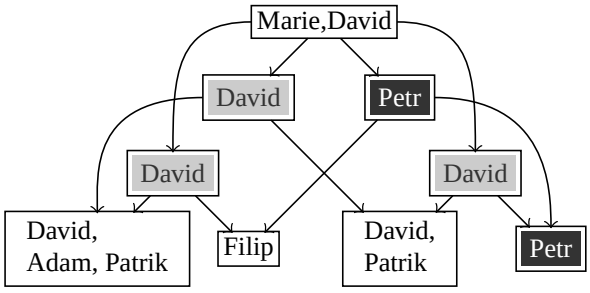


Figure 1: $\{\mathcal{O}_u(t_i\!\uparrow) \mid t_i \in R^*\} \subseteq \mathscr{I}(Q)$ with a relation $\sqsubseteq$

Note that $\mathcal{O}_u$ generally is not an injection. In other words, $\mathcal{O}_u(t_i\!\uparrow) = \mathcal{O}_u(t_j\!\uparrow)$ for some $t_i\!\uparrow \neq t_j\!\uparrow$. In particular, $\mathcal{O}_u(t_i\!\uparrow) = \mathcal{O}_u(t_j\!\uparrow) =$ "Petr" and $\mathcal{O}_u(t_k\!\uparrow) = \mathcal{O}_u(t_l\!\uparrow) = \mathcal{O}_u(t_m\!\uparrow) =$ "David". To get an ordering, we need to resolve these duplicities:

- First, as the occurrences of "Petr" are in the relation $\sqsubseteq$, we drop the less "preferred" one.
- In the case of the triplet of occurrences of "David", we are unable to determine the one with the highest "preference". Nevertheless, notice that:
  - The set $\{$Marie, David$\}$ is preferred to any of the occurrences of "David". In other words, whichever the most preferred occurrence of "David" is, it is less preferred then the set $\{$Marie, David$\}$.

– There is a unique occurrence of "Filip", for which we can find an occurrence of "David" with a higher preference. In other words, whichever the most preferred occurrence of "David" is, it is preferred more then the occurrence of "Filip". The same rationale applies for the sets {David, Adam, Patrik} and {David, Patrik}.

Thus, we get the resulting order, depicted in the following figure:

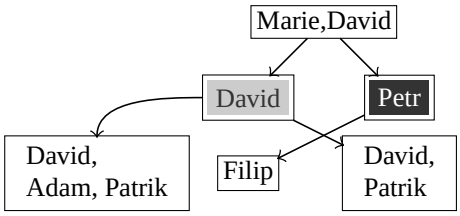


Figure 2: Ordering $\preceq^{\mathscr{I}(Q)}$ on $\{\mathcal{O}_u(t_i\uparrow) \mid t_i \in R^*\} \subseteq \mathscr{I}(Q)$ □

To sum up, the order $\preceq^{\mathscr{I}(Q)}$ on the resulting collection $\{\mathcal{O}_u(t_i\uparrow) \mid t_i \in R^*\} \subseteq \mathscr{I}(Q)$ of sets is defined as:

$$\mathcal{O}_u(t_i\uparrow) \preceq^{\mathscr{I}(Q)} \mathcal{O}_u(t_j\uparrow) \Leftrightarrow \\ (\forall t_k \in R^*)\big([\mathcal{O}_u(t_k\uparrow) = \mathcal{O}_u(t_i\uparrow)] \Rightarrow (\exists t_l \in R^*)[\mathcal{O}_u(t_l\uparrow) = \mathcal{O}_u(t_j\uparrow) \ \wedge \ t_k \preceq^{R^*} t_l]\big)$$

As the minimal set of relational algebra operations consists of two unary (restriction and projection) and three binary operations (set difference, set union, and cartesian product), a binary operations need also to be considered. In general, a binary operation

$$\mathcal{O}_b : \mathscr{I}(R) \times \mathscr{I}(R') \to \mathscr{I}(Q)$$

is a mapping from a couple of sets of relational instances of R and R' into a set of relational instances of resulting relation Q. As in the foregoing example, we get a resulting collection

$$\{\mathcal{O}_b(t_i\uparrow, t'_k\uparrow) \mid (t_i, t'_k) \in R^* \times R'^*\} \subseteq \mathscr{I}(Q)$$

of sets, and the order $\preceq^{\mathscr{I}(Q)}$ definition:

$$\mathcal{O}_b(t_i\uparrow, t'_k\uparrow) \preceq^{\mathscr{I}(Q)} \mathcal{O}_b(t_j\uparrow, t'_l\uparrow) \Leftrightarrow \\ [\forall (t_m, t'_p) \in R^* \times R'^*]\big([\mathcal{O}_b(t_m\uparrow, t'_p\uparrow) = \mathcal{O}_b(t_i\uparrow, t'_k\uparrow)] \Rightarrow \\ [\exists (t_n, t'_q) \in R^* \times R'^*][\mathcal{O}_b(t_n\uparrow, t'_q\uparrow) = \mathcal{O}_b(t_j\uparrow, t'_l\uparrow) \ \wedge \ t_m \preceq^{R^*} t_n \ \wedge \ t'_p \preceq^{R'^*} t'_q]\big)$$

What are the consequences of this approach? Generally

$$\mathcal{O}_b(t_i\uparrow, t'_k\uparrow) \preceq^{\mathscr{I}(Q)} \mathcal{O}_b(t_j\uparrow, t'_l\uparrow) \Rightarrow \mathcal{O}_b(t_i\uparrow, t'_k\uparrow) \supseteq \mathcal{O}_b(t_j\uparrow, t'_l\uparrow)$$

does not hold for nonmonotonic operations (consider the relational operation of difference). With respect to the relational property of *closure*, it is clear that the framework for defining preference on the tuples of relational instances needs to be generalized. We need to express the preference structure on the powerset $\mathscr{I}(R)$ of all possible instances R^* of a relation R.

The foregoing definitions of orderings on sets correspond to Hoare's ordering. It is one of the orderings provided by the study of semantics of non-determinism [26, 19]. Possibility of employing another semantics, however, needs further investigation.

3. Summary and Future Work

A framework for relational algebra with preferences is proposed. It should present a basis for semantically rich, easy to handle and flexible preference model aiming at deep personalization of database queries. It differs from the other above mentioned approaches in that the preferences are not part of the queries but part of relational instances. Is is shown that they need to be defined between sets of elements – relational instances. This approach is strictly more general allowing to model nonmonotonic operations.

It can also be shown that the proposed approach generalizes the fuzzy sets-based approach to modelling uncertainty in database systems. The reason is that there exists a homomorphism (A and B stand for sets of attributes of relations R):

$$\Big(\big\{\mathscr{I}_F\big(R(A\cup B)\big)\cup\mathscr{I}_F\big(R(A)\big)\cup\mathscr{I}_F\big(R(B)\big)\big\};\Omega_{\mathcal{F}}\Big)\longrightarrow\Big(\big\{\mathscr{I}_O\big(R(A\cup B)\big)\cup\mathscr{I}_O\big(R(A)\big)\cup\mathscr{I}_O\big(R(B)\big)\big\};\Omega_O\Big)$$

of the algebra consisting of

- the support comprising all the fuzzy relation instances that can arise by means of monotonic fuzzy relational algebra operations applied on relations $R(A\cup B), R(A), R(B)$, and of
- monotonic fuzzy relational algebra operations $\Omega_{\mathcal{F}}$

into the algebra consisting of

- the support comprising all the relation instances with ordering that can arise and whose ordering is defined by means of monotonic preference algebra operations applied on relations $R(A\cup B), R(A), R(B)$, and of
- monotonic preference relational algebra operations Ω_O

Moreover, this homomorphism is unique for all t-norms.

Furthermore, it can be shown that associativity and commutativity of the original union, product, restrict, and project operations are retained. Specifically for the generalized restrict operation, the following equivalences, which hold for the classical restrict operation, are retained:

$$\begin{aligned}\sigma_{\varphi_1\vee\varphi_2}(R) &\equiv \sigma_{\varphi_1}(R)\cup\sigma_{\varphi_2}(R)\\ \sigma_{\varphi_1\wedge\varphi_2}(R) &\equiv \sigma_{\varphi_1}(R)\cap\sigma_{\varphi_2}(R)\\ \sigma_{\neg\varphi}(R) &\equiv R-\sigma_{\varphi}(R)\end{aligned}$$

Using the proposed approach, other relational operations (intersect, join, and divide), also, retain the usual properties of their classical relational counterparts:

$$\begin{aligned}R\cap S &\equiv R-(R-S)\\ R\div S &\equiv \pi_{A-B}(R)-\pi_{A-B}\Big(\big(\pi_{A-B}(R)\times S\big)-R\Big)\\ R\bowtie S &\equiv \pi_{A\cup B}\big(\sigma_{\varphi}(R\times S)\big)\end{aligned}$$

These results are promising for many query optimization issues, which present many open problems. In particular, evaluation and optimization of preference queries, including cost-base optimization need to be addressed.

Another open problem is how the proposed framework fits into the relational data model. A promising approach seems to be an array-based data model proposed by [22, 23]. The key component of this model

is an ordered data structure called an *arrable*, for array-table. Informally, it is a collection of named arrays that, in their simplest form, are vectors of elements of base type. In this form, an arrable is essentially a table organized by columns. Among others, arrable facilitates the implementation of algebraic operators by means of *realizers* [18] – a set of linear extensions uniquely characterizing a given partial order. Realizers offer an effective way to implement relational operations with ordering. Nevertheless, a comparison with other possibly effective implementations of relational operations needs further investigation.

List of symbols

$\Pi(R^*)$ $\{A \mid A$ is a fuzzy subset of $R^*\}$,

$\mathscr{I}(R)$ set of all possible instances R^* of a relation R

$\mathscr{I}_F(R)$ set of all possible fuzzy instances R^{F*} of a relation R

$\mathscr{I}_O(R)$ set of all possible ordered instances of a relation R

References

[1] M. Lacroix and P. Lavency, "Preferences; Putting More Knowledge into Queries.," in *VLDB* (P. M. Stocker, W. Kent, and P. Hammersley, eds.), pp. 217–225, Morgan Kaufmann, 1987.

[2] W. Kießling, "Foundations of Preferences in Database Systems," in *Proceedings of the 28th VLDB Conference*, (Hong Kong, China), pp. 311–322, 2002.

[3] W. Kießling, "Optimatization of Relational Preference Queries," in *Conferences in Research and Practice in Information Technology* (H. Williams and G. Dobbie, eds.), vol. 39, (University of Newcastle, Newcastle, Australia), Australian Computer Society, 2005.

[4] W. Kießling, "Preference Queries with SV-Semantics.," in *COMAD* (J. Haritsa and T. Vijayaraman, eds.), pp. 15–26, Computer Society of India, 2005.

[5] W. Kießling and B. Hafenrichter, "Algebraic optimization of relational preference queries," Tech. Rep. 2003-01, Institute of Computer Science, University of Augsburg, February 2003.

[6] W. Kießling, "Preference constructors for deeply personalized database queries," Tech. Rep. 2004-07, Institute of Computer Science, University of Augsburg, March 2004.

[7] B. Hafenrichter and W. Kießling, "Optimization of relational preference queries," in *CRPIT '39: Proceedings of the sixteenth Australasian conference on Database technologies*, (Darlinghurst, Australia, Australia), pp. 175–184, Australian Computer Society, Inc., 2005.

[8] J. Chomicki, "Preference Formulas in Relational Queries," *ACM Trans. Database Syst.*, vol. 28, no. 4, pp. 427–466, 2003.

[9] J. Chomicki, "Semantic optimization techniques for preference queries," 2005.

[10] J. Chomicki and J. Song, "Monotonic and nonmonotonic preference revision," 2005.

[11] J. Chomicki, S. Staworko, and J. Marcinkowsk, "Preference-driven querying of inconsistent relational databases," in *Proc. International Workshop on Inconsistency and Incompleteness in Databases*, (Munich, Germany), March 2006.

[12] S. Börzsönyi, D. Kossmann, and K. Stocker, "The skyline operator," in *Proceedings of the 17th International Conference on Data Engineering*, (Washington, DC, USA), pp. 421–430, IEEE Computer Society, 2001.

[13] R. Agrawal and E. Wimmers, "A Framework for Expressing and Combining Preferences.," in *SIGMOD Conference* (W. Chen, J. F. Naughton, and P. A. Bernstein, eds.), pp. 297–306, ACM, 2000.

[14] V. Hristidis, N. Koudas, and Y. Papakonstantinou, "Prefer: a system for the efficient execution of multi-parametric ranked queries," in *SIGMOD '01: Proceedings of the 2001 ACM SIGMOD international conference on Management of data*, (New York, NY, USA), pp. 259–270, ACM Press, 2001.

[15] W. Ng, "An Extension of the Relational Data Model to Incorporate Ordered Domains," *ACM Transactions on Database Systems*, vol. 26, pp. 344–383, September 2001.

[16] S. Raghavan and H. Garcia-Molina, "Complex queries over web repositories," 2003.

[17] R. Nedbal, "Relational Databases with Ordered Relations," *Logic Journal of the IGPL*, vol. 13, no. 5, pp. 587–597, 2005.

[18] D. R. Raymond, *Partial-order databases*. PhD thesis, University of Waterloo, Waterloo, Ontario, Canada, 1996. Adviser-W. M. Tompa.

[19] L. Libkin, *Aspects of partial information in databases*. PhD thesis, University of Pensylvania, Philadelphia, PA, USA, 1995.

[20] P. Seshadri, M. Livny, and R. Ramakrishnan, "The design and implementation of a sequence database system," in *VLDB '96: Proceedings of the 22th International Conference on Very Large Data Bases*, (San Francisco, CA, USA), pp. 99–110, Morgan Kaufmann Publishers Inc., 1996.

[21] R. Ramakrishnan, D. Donjerkovic, A. Ranganathan, K. S. Beyer, and M. Krishnaprasad, "Srql: Sorted relational query language," in *SSDBM '98: Proceedings of the 10th International Conference on Scientific and Statistical Database Management*, (Washington, DC, USA), pp. 84–95, IEEE Computer Society, 1998.

[22] A. Lerner and D. Shasha, "Aquery: Query language for ordered data, optimization techniques, and experiments," in *29th International Conference on Very Large Data Bases (VLDB'03)*, (Berlin, Germany), pp. 345–356, Morgan Kaufmann Publishers, September 2003.

[23] A. Lerner, *Querying Ordered Databases with AQuery*. PhD thesis, ENST-Paris, France, 2003.

[24] V. Subrahmanian, *Uncertainty in Databases and Knowledge Bases*, pp. 315–411. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1997.

[25] R. Nedbal, "Fuzzy database systems – concepts and implementation," Master's thesis, Czech Technical University, Faculty of Nuclear Sciences and Physical Engineering, Prague, June 2003.

[26] P. Buneman, A. Jung, and A. Ohori, "Using powerdomains to generalize relational databases," *Theor. Comput. Sci.*, vol. 91, no. 1, pp. 23–55, 1991.

Vyhledávání vědeckých lékařských informací v databázi MEDLINE/PubMed se zaměřením na potřeby medicíny založené na důkazech (evidence-based medicine)

doktorand:
MUDR. VENDULA PAPÍKOVÁ
Oddělení medicínské informatiky
Ústav informatiky AV ČR
Pod Vodárenskou věží 2
182 07 Praha 8
papikova@euromise.cz

školitel:
DOC. PHDR. RUDOLF VLASÁK
Ústav informačních studií a knihovnictví
Filozofická fakulta Univerzity Karlovy
U Kříže 8
158 00 Praha 5
rudolf.vlasak@ff.cuni.cz

obor studia:
Informační věda

Práce byla částečně podpořena výzkumným záměrem AV0Z10300504.

Abstrakt

Cílem práce bylo vybrat veřejně dostupné nástroje umožňující efektivní vyhledávání publikací v databázi MEDLINE/PubMed s ohledem na specifické potřeby medicíny založené na důkazech. Detailně jsou popsány následující systémy: PubMed Clinical Queries, PICO, AskMEDLINE, PubMed Interact a PubMatrix.

Klíčová slova: evidence-based medicine, EBM, medicína založená na důkazech, vědecké lékařské informace, metody vyhledávání, řešeršní systémy, MEDLINE/PubMed, PubMed Clinical Queries, PICO, AskMEDLINE, PubMed Interact, SLIM, alternativní interface, PubMatrix

1. Úvod

Medicína založená na důkazech (evidence-based medicine, EBM) je mezioborová, klinicky orientovaná disciplína, která se systematicky rozvíjí od počátku 90. let 20. století (její základy však byly položeny mnohem dříve). V lékařské literatuře se termín evidence-based medicine poprvé objevil v roce 1991 (Guyatt a kol.) [1], [2]. Její obsah a cíle vymezuje dnes již klasická definice: *Medicína založená na důkazech je svědomité, jednoznačné a kritické uplatňování nejnovějších a nejlepších důkazů při rozhodování o péči o jednotlivé pacienty. Vykonáváním EBM v praxi je myšlena integrace individuální klinické odbornosti lékařů s nejkvalitnějšími objektivními důkazy pocházejícími ze systematicky prováděného výzkumu* [3].
Za duchovního otce EBM je považován skotský epidemiolog profesor Archie Cochrane (1908–1988), díky jehož úsilí byl koncept EBM postupně přijat. Jeho práce byly podnětem pro vznik kontrolovaných klinických studií a pro systematické zpracovávání klinických a epidemiologických dat do metaanalýz, tzv. systematických přehledů (**Cochrane reviews, systematic reviews**). Cochranovy systematické přehledy jsou vysoce kvalitní zdroje informací o určitém tématu nebo klinické otázce. Vznikají v metodicky přesně definovaném a reprodukovatelném procesu, jehož součástí je velmi pečlivé a důkladné vyhledávání primárních vědeckých dokumentů (publikovaných i nepublikovaných) a kritické posouzení jejich validity (k dalšímu zpracování jsou vybrány pouze studie odpovídající stanoveným kritériím). Cílem tohoto procesu je minimalizovat riziko systematické chyby (bias) a získat tak co možná nejspolehlivější závěry [4], [6].
Cochranovy systematické přehledy jsou alternativou k tradičnímu pojetí přehledných článků (reviews), které bývají zatíženy velkou nepřesností, nemusí nutně reprezentovat nejlepší a nejnovější informace potřebné pro správné rozhodování a s ohledem na nároky medicíny založené na důkazech mohou být zastaralé. Průběžná aktualizace systematických přehledů Cochranova typu je prováděna čtvrtletně. V úvodu každého takového přehledu najdeme datum posledního prohledávání informačních zdrojů a datum poslední podstatné provedené změny [5].
Medicína založená na důkazech klasifikuje různé typy klinických důkazů podle stupně jejich nezávislosti na

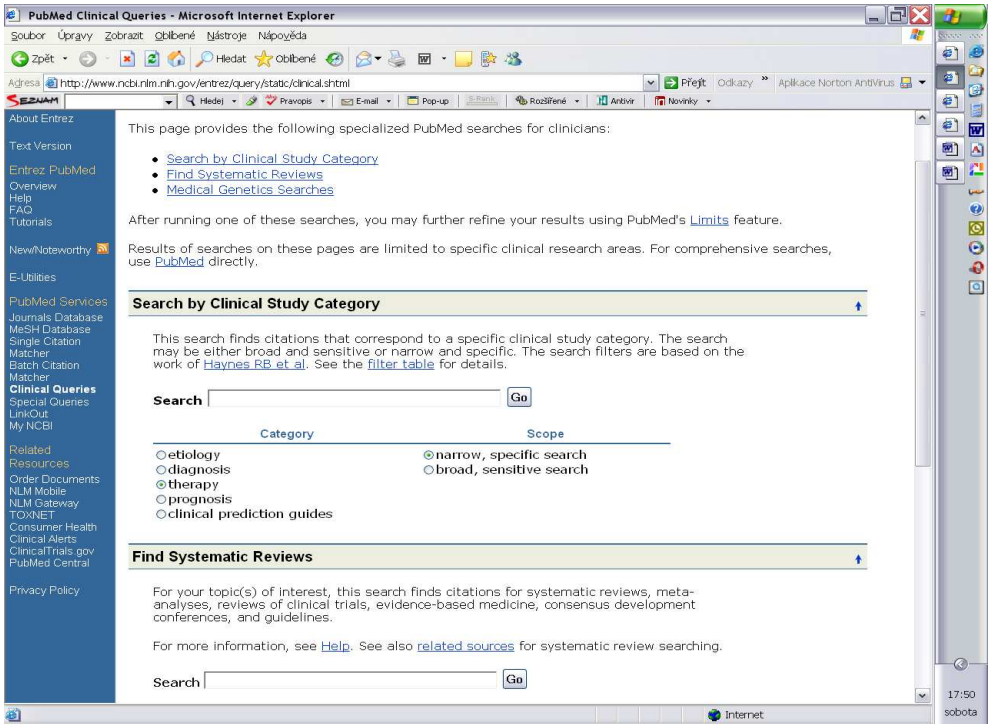
vlivech způsobujících nepřesnosti a systematické chyby. Důkazy s nejvyšší vahou pro terapeutické postupy poskytují **randomizované, dvojitě slepé, placebo kontrolované studie**. Naopak pacientova svědectví, kazuistiky či dokonce názory odborníků mají podle pravidel EBM malý význam.

2. Cíl a metodika

Cílem práce bylo vybrat mezi nástroji umožňujícími vyhledávání vědeckých lékařských informací v databázi MEDLINE takové, které jsou užitečné s ohledem na potřeby a kritéria medicíny založené na důkazech.

Má-li být medicína založená na důkazech praktikována širokou lékařskou veřejností, je nutné, aby přístup k informacím nebyl licenčně limitován. Z tohoto důvodu byly pro tuto práci vybrány nástroje, které jsou volně dostupné prostřednictvím internetu a jako zdroj používají databázi MEDLINE veřejně přístupnou přes webové rozhraní PubMed při Národní lékařské knihovně USA v Bethesdě (`http://www.pubmed.org`).

Systémů určených k vyhledávání literatury v databázi MEDLINE/PubMed je dnes k dispozici celá řada. Jejich rozvoj v posledních letech akceleroval především ohromný nárůst vědeckých publikací na jedné straně a současně specifické potřeby uživatelů (např. vědeckých pracovníků, bioinformatiků, klinických lékařů) na straně druhé.
Nástroje popisované v tomto článku spojuje skutečnost, že obsahují funkce ”šité na míru” klinicky zaměřené problematice EBM. Jedná se o filtry PubMed Clinical Queries, vyhledávače PICO a AskMEDLINE, webové rozhraní SLIM a jeho pokročilejší verzi PubMed Interact. Navíc byl otestován software PubMatrix, který je primárně komponován spíše jako bioinformatický nástroj, jeho některé přednosti však mohou být úspěšně využity také pro účely medicíny založené na důkazech.



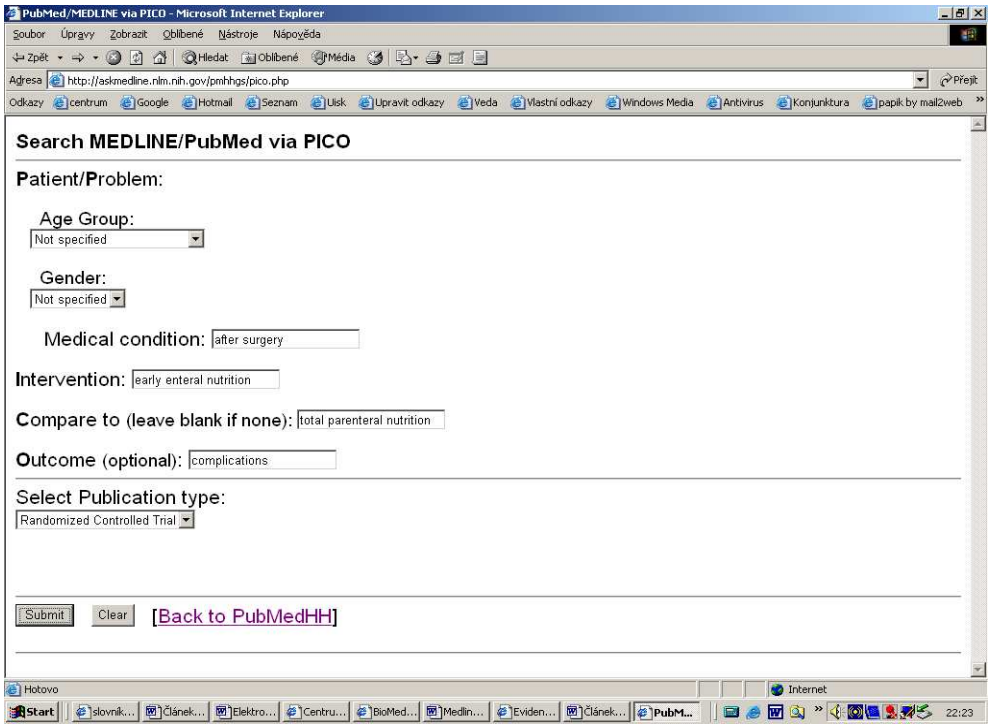
Obrázek 1: Formulář pro vyhledávání pomocí PubMed Clinical Queries

PubMed Clinical Queries

Jednu z možností, jak efektivně vyhledat v databázi MEDLINE/PubMed klinicky významné publikace, najdeme přímo v rámci rozhraní PubMed. Nástroj k tomu určený se jmenuje Clinical Queries a cesta k němu vede přes hypertextový odkaz v levé nástrojové liště hlavní strany rozhraní PubMed (`http://www.pubmed.org`).
Pomocí Clinical Queries je možné vyhledávat jednak studie podle klinických kategorií (**Clinical Study Category**: etiologie, diagnóza, terapie, prognóza a návody pro klinické předpovědi), jednak systematické přehledy (**Systematic Reviews**), (obr. 1). Kromě pravých systematických přehledů Cochranova typu tento filtr vyhledává navíc také metaanalýzy, přehledy klinických studií, termín evidence-based medicine, konference formulující shodná stanoviska a praktická doporučení (guidelines). Širší pojetí termínu ”systematic review” na první pohled vnáší riziko vzniku nepřesností a matení pojmů, v rešeršní praxi se ale jeví jako výhodné z toho důvodu, že pravých systematických přehledů Cochranova typu je stále ještě nedostatek a jejich vyhledávání v řadě medicínských oborů by poskytovalo nulové výsledky.

PICO

PICO je vyhledávací rozhraní, které pomocí speciálně přizpůsobeného formuláře vede uživatele k sestavení správně formulované otázky podle doporučení evidence-based medicine [7]. Do jednotlivých vyhledávacích polí uživatel postupně vkládá klíčová slova podle vzorce PICO: zadané termíny mají postupně charakterizovat: pacienta/problém (**Patient/Problem**), léčebný postup (**Intervention**), srovnávaný léčebný postup (**Comparison**) a očekávaný výsledek (**Outcome**) [8], [9], (obr. 2). Je možné definovat také požadovaný typ dokumentu - rozbalovací menu v dolní části rozhraní nabízí pět typů dokumentů vhodných pro EBM praxi (klinické studie, metaanalýzy, randomizované kontrolované studie, přehledy a praktická doporučení). Vyhledané citace zahrnují odkaz na abstrakt, plný text a související články.



Obrázek 2: Vyhledávací rozhraní PICO

Vývoj vyhledávače PICO byl iniciován potřebami zaneprázdněných lékařů, kteří mají zájem o aplikaci medicíny založené na důkazech v klinické praxi, ale nejsou seznámeni s kontrolovanými slovníky a vyhledávacími technikami umožňujícími efektivní vyhledávání vědecké lékařské literatury [6].

Metodou PICO je možné vyhledávat pomocí rozhraní na internetových stránkách Národní lékařské knihovny v Bethesdě (odkaz na vyhledávací formulář je součástí rozhraní "PubMed for Handhelds" (`http://pubmedhh.nlm.nih.gov/nlm`) nebo na adrese `http://askmedline.nlm.nih.gov/ask/pico.php`, kde je navíc zabudován "spelling checker" kontrolující pravopisné chyby a překlepy.

Případová úloha 1:

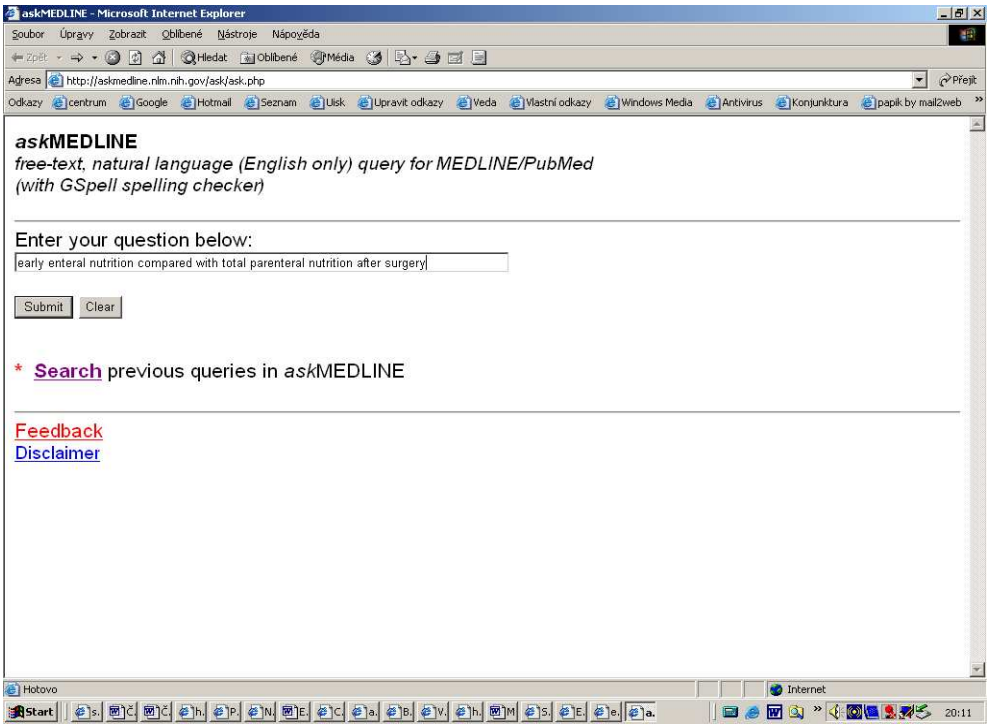
Zadání: Je časné zavedení enterální výživy u pacientů po operaci spojeno s menším rizikem pooperačních komplikací ve srovnání s pacienty s úplnou parenterální výživou?

P	I	C	O
pacienti po operaci	časná enterální výživa	úplná parenterální výživa	pooperační komplikace

Výsledek: Bylo nalezeno celkem 20 dokumentů, z toho 8 randomizovaných kontrolovaných studií a dva přehledové články.

AskMEDLINE

Asi nejjednodušší způsob, jak získat relevantní citace článků zabývající se klinicky orientovanou problematikou, nabízí nástroj zvaný AskMEDLINE (`http://askmedline.nlm.nih.gov`). Jedná se o rozhraní, které vzniklo na podkladě vyhledávače PICO (viz výše). Jeho tvůrci se pokusili zjednodušit klinickým lékařům vyhledávání v databázi MEDLINE/PubMed v maximální možné míře pomocí metod založených na přirozeném jazyku. AskMEDLINE je nástroj pro jednoduché jazykové dotazování, který umožňuje uživatelům prohledávat databázi MEDLINE/PubMed pomocí přirozeně formulované otázky (**natural language query**). Uživatel pouze formuluje dotaz nebo problém do vyhledávacího řádku, aniž by musel znát pravidla pro efektivní vyhledávání v databázi MEDLINE/PubMed (obr. 3). AskMEDLINE transformuje zadání dotazu podle pravidel PICO a nabídne seznam citací spolu s odkazy na abstrakt, plný text a související články [10]. V případě, že uživatel není s výsledkem rešerše spokojen, nabídne tento vyhledávač odkaz přímo na formulář PICO.



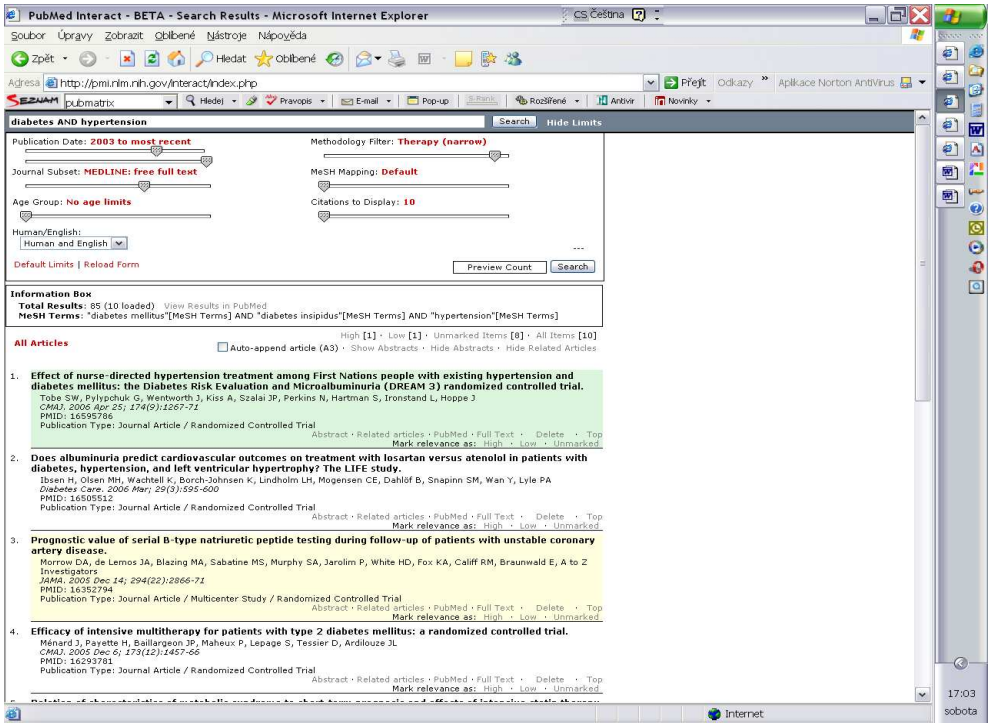
Obrázek 3: Vyhledávací rozhraní AskMEDLINE

Případová úloha 2:

Zadání: Co je lepší pro pacienty po operaci - časná enterální výživa nebo totální parenterální výživa?
Výsledek: Bylo nalezeno 59 klinicky relevantních dokumentů.

PubMed Interact

PubMed Interact je alternativní vyhledávací rozhraní pro MEDLINE/PubMed s intuitivními a interaktivními prvky (obr. 4). Na rozdíl od klasického rozhraní (dostupného na `http://www.pubmed.org`), které pro efektivní vyhledávání textových dokumentů vyžaduje relativně dost znalostí, zkušeností a času, PubMed Interact významně tyto nároky snižuje a navíc umožňuje řadu pokročilých funkcí.
Jedná se o dokonalejší verzi rozhraní zvaného **SLIM (Slider Interface for MEDLINE/PubMed Searches** [11], `http://pmi.nlm.nih.gov/slide`), která je zatím stále testována (beta-verze) a technicky vylepšována. Nicméně už nyní je plně funkční na adrese `http://pmi.nlm.nih.gov/interact` [12].



Obrázek 4: Vyhledávací formulář v rozhraní PubMed Interact se záznamem části rešerše

PubMed Interact [13] využívá ve vyhledávacím formuláři šest posuvných jezdců (”slider bars”) pro snadné a rychlé definování následujících limitů: časové vymezení (Publication Date), podskupiny časopisů (Journal Subset), věkové skupiny (Age Group), vymezení typu článku, resp. použité metodologie (Metodology Filter) a mapování MeSH terminologie (Mesh Mapping). Limity pro vyhledávání lze jednoduše zadat uchopením posuvného jezdce na monitoru pomocí počítačové myši a tažením doprava nebo doleva. Při pohybu jezdce se nad ním zobrazují textová pole označující konkrétní hodnotu limitu. Souhrnně lze říci, že díky možnosti nastavovat vyhledávací limity pomocí posuvných jezdců **PubMed Interact velmi usnadňuje možnost redefinovat a přesměrovat vyhledávací strategie v průběhu rešerše s ohledem na aktuálně získané výsledky**.
Pro potřeby evidence-based medicine je velice užitečná především možnost snadného nastavení metodologického filtru. Filtr zahrnuje tři hlavní kategorie publikací: kazuistiky (case reports), klinické studie (clinical study categories) a systematické přehledy (systematic reviews). Kategorie klinických studií obsahuje dalších deset variant nastavení filtrů pro vyhledávání v rámci klinických studií. Kategorie systematických review je, stejně jako v PubMed Clinical Queries (viz výše), pojata poněkud šířeji, než praví v úvodu cito-

vaná definice systematických přehledů Cochranova typu.

Vyhledávač pracuje s Booleovskými operátory, rozšíření slovního základu (truncation) a pokročilé vyhledávání pomocí definování prohledávaných polí (autor, titul/abstrakt atd.) známé z klasického rozhraní PubMed je rovněž možné, musí se však zadat manuálně do příkazového řádku.

Rozbalovací menu v levém dolním rohu vyhledávacího formuláře PubMed Interact umožňuje také nastavení vyhledávání publikací pouze v anglickém jazyce a dále publikací zabývající se problematikou pouze lidské populace.

Pro praxi velmi užitečná je také funkce zvaná "Preview Results Count" umožňující jediným kliknutím zjistit, kolik záznamů bude zadaná rešerše obsahovat. Stejně jako systém posuvných jezdců i ona zkracuje dobu potřebnou k nalezení záznamů s co možná nejvyšší relevancí na minimum. V případě nevyhovujícího počtu záznamů (ať už velmi vysokého nebo velmi nízkého) je možné upravit limity pro vyhledávání ihned a nečekat na načítání stránky, která nebude vyhovovat uživatelovým potřebám.

Kromě výše uvedených vyhledávacích limitů popisované rozhraní umožňuje nastavit limity pro zobrazení výsledků na monitoru. Jde o volbu počtu zobrazených záznamů (Citations to Display) pomocí posuvného jezdce a dále o přepínač v pravém horním rohu vyhledávacího formuláře, který umožňuje uvolnění prostoru na obrazovce pro pohodlnější prohlížení výsledků rešerše tím, že vypne zobrazení limitů pro vyhledávání (Hide/Show Limits).

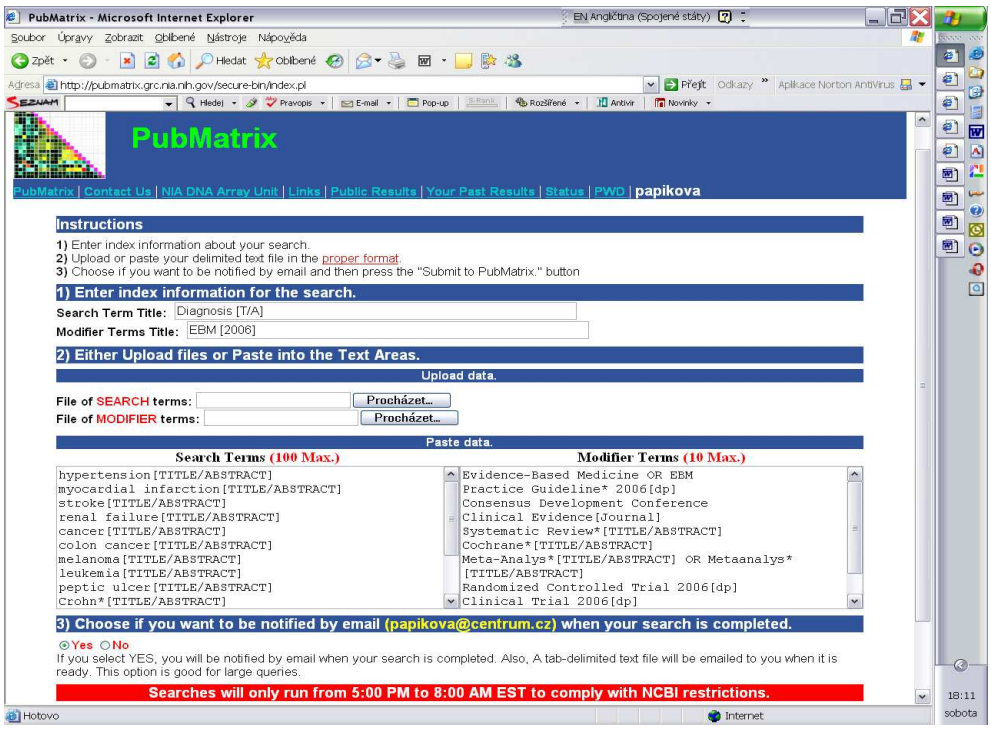
Z uživatelského hlediska jsou velmi cenné i další funkce systému, které umožňují doplnit a upravit seznam vyhledaných citací tak, aby zcela vyhovoval informační potřebě uživatele, přičemž zobrazení všech provedených akcí zůstává na téže stránce. Práce s citacemi je díky tomu přehledná a uživatel se neztrácí v "odkazech na odkazy". Konkrétně se jedná o tyto možnosti:

- Abstrakta jsou strukturovaná a lze je zobrazit na téže stránce přímo pod příslušnou citací, a to buď jednotlivě nebo u všech článků najednou.
- Odkazy na související články se zobrazují rovněž na téže stránce pod odpovídající citací. Do každého ze souvisejících článků je možné nahlédnout a poté ho buď zavřít nebo přidat k rešerši pod příslušnou citaci (pořadí citací se přitom automaticky přečísluje). Také zmiňovaný náhled souvisejícího článku se zobrazuje na stejné stránce, pro zachování přehlednosti je umístěn ve žlutě označeném poli.
- Systém umožňuje vymazání citací s nízkou relevancí a barevné označení vyhovujících citací dle stupně relevance (vysoká relevance - zeleně, nízká relevance - žlutě, průměrně relevantní zůstávají bílé). Každou takto vzniklou skupinu je možné zobrazit jako zvláštní seznam.

Všechny volby v PubMed Interact jsou sestrojeny tak, aby jejich nastavení bylo možné jediným kliknutím nebo tahem počítačovou myší. Díky tomu je práce s tímto rozhraním velmi pohodlná a svižná. Toto spolu s dalšími atributy kvalitního vyhledávače z něj dělá velmi zajímavý a užitečný nástroj pro online vyhledávání lékařské literatury pro potřeby EBM uplatňované v praxi.

PubMatrix

PubMatrix je nástroj volně dostupný na internetu (`http://pubmatrix.grc.nia.nih.gov`) pro jednoduché vytěžování vědeckých textů spravovaných v databázi MEDLINE/PubMed s použitím jakýchkoliv dvou seznamů klíčových slov. Provádí mnohočetné srovnávání studovaných termínů, jehož výsledkem je frekvenční matice současného výskytu vždy dvou zpárovaných termínů (frequency matrix of term co-occurence). Číselné údaje v matici jsou propojeny s odpovídajícími rešeršemi v databázi MEDLINE/PubMed hypertextovými odkazy. Díky možnosti srovnávat velké seznamy klíčových slov PubMatrix významně snižuje čas potřebný k získání hledaných informací, přičemž nalezené dokumenty předkládá systematicky roztříděné ve výstupní matici [14].
S ohledem na uvedené je PubMatrix velmi cenným nástrojem pro vytěžování velkých souborů molekulárně



Obrázek 5: PubMatrix - dotazovací formulář

genetických dat a pro tyto účely byl také primárně navržen. V předkládané práci je popsána zcela odlišná aplikace. Pomocí deseti vybraných termínů úzce spojených s problematikou EBM byla provedena de facto **mnohočetná rešerše** mapující tuto oblast z hlediska vybraných diagnóz (obr. 5, 6).

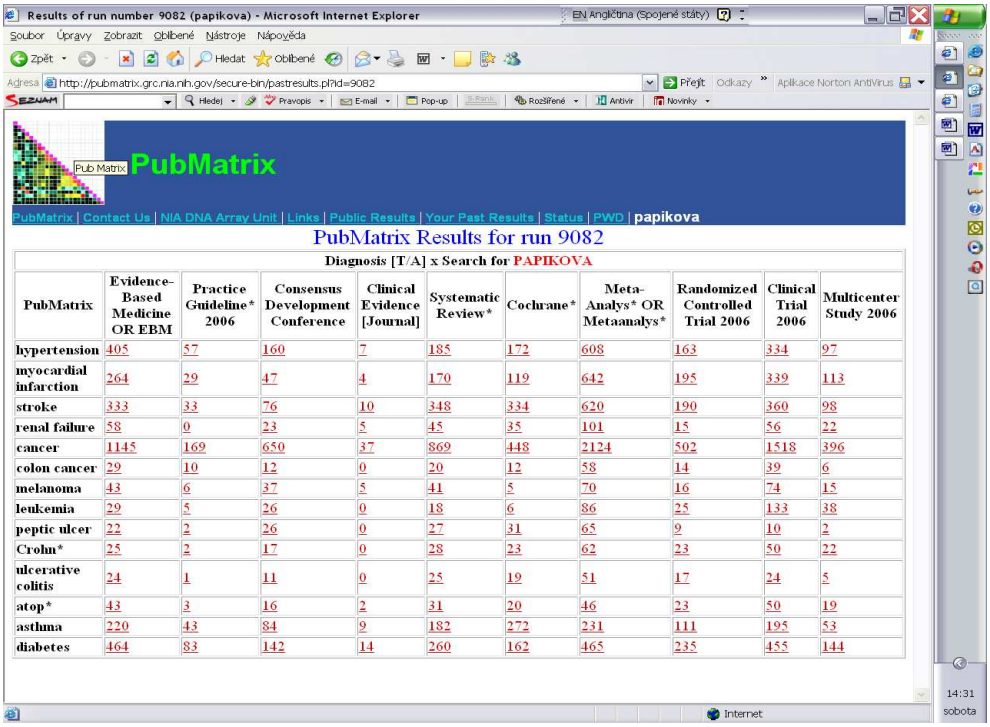
Postup pro vyhledávání informací dle zadané případové úlohy pomocí PubMatrix je následující:

1. Po povinné (bezplatné) registraci se otevře dotazovací stránka pro vepsání vybraných klíčových slov do dvou sloupců: ”search terms” (maximálně 100) a ”modifier terms” (maximálně 10).
2. Zadání dvou seznamů klíčových slov s použitím pravidel pro popis a syntaxi platných v rozhraní PubMed (obr. 5).
3. Volba zaslání výsledků e-mailem (během testování pro účely této publikace nebyla funkční).
4. Odeslání zadání.

Výsledkem provedené případové úlohy je matice umožňující okamžité prohlížení 140 rešerší (obr. 6).

Výhody systému:

- *Jednoduchost*: Rozhraní dostupné prostřednictvím internetu nevyžaduje složitou průpravu a počet kroků k zadání dotazu je minimální (viz postup).
- *Časová úspornost*: Po zadání jediného vyhledávacího formuláře získáme přístup k několika (nebo dokonce k mnoha) nezávislým rešerším. Jejich otevření je možné jediným kliknutím na kterékoliv pole výsledné matice.
- *Náhled na studovanou problematiku z hlediska četnosti nalezených dokumentů obsahujících hledané dvojice termínů*. Taková vlastnost vyhledávacího systému umožňuje rychlou orientaci v neznámé



Results of run number 9082 (papikova) - Microsoft Internet Explorer

PubMatrix

PubMatrix | Contact Us | NIA DNA Array Unit | Links | Public Results | Your Past Results | Status | PWD | papikova

PubMatrix Results for run 9082

Diagnosis [T/A] x Search for PAPIKOVA

PubMatrix	Evidence-Based Medicine OR EBM	Practice Guideline* 2006	Consensus Development Conference	Clinical Evidence [Journal]	Systematic Review*	Cochrane*	Meta-Analys* OR Metaanalys*	Randomized Controlled Trial 2006	Clinical Trial 2006	Multicenter Study 2006
hypertension	405	57	160	7	185	172	608	163	334	97
myocardial infarction	264	29	47	4	170	119	642	195	339	113
stroke	333	33	76	10	348	334	620	190	360	98
renal failure	58	0	23	5	45	35	101	15	56	22
cancer	1145	169	650	37	869	448	2124	502	1518	396
colon cancer	29	10	12	0	20	12	58	14	39	6
melanoma	43	6	37	5	41	5	70	16	74	15
leukemia	29	5	26	0	18	6	86	25	133	38
peptic ulcer	22	2	26	0	27	31	65	9	10	2
Crohn*	25	2	17	0	28	23	62	23	50	22
ulcerative colitis	24	1	11	0	25	19	51	17	24	5
atop*	43	3	16	2	31	20	46	23	50	19
asthma	220	43	84	9	182	272	231	111	195	53
diabetes	464	83	142	14	260	162	465	235	455	144

Obrázek 6: PubMatrix - výsledná matice

oblasti, zmapování oboru, vytipování hojně studovaných nebo naopak zcela nepopsaných vztahů a nalezení nových souvislostí nebo překvapivých skutečností.

- *Rešerše je možné prohlížet opakovaně* (výsledná matice zůstává uložena na internetu), celou matici je možné navíc *aktualizovat* (opět jediným kliknutím počítačové myši).

Limity systému:

- Počet termínů, které je možné zadat do veřejně dostupné verze software je omezen na 10 x 100.

- Maximální počet dokumentů, které systém přiřadí k jednotlivým polím matice je 100 000. Z praktického rešeršního hlediska to však není významné (rešerše o tak velkém počtu záznamů přestává být rešerší).

- Systém je přístupný v omezeném čase (5 PM EST - 8 AM EST, tj. v ČR od 23:00 do 14:00). Zadání klíčových slov do vyhledávacího formuláře je však možné provést kdykoliv a pro výsledky se vrátit později (přístup k již hotovým maticím není časově limitován).

- V případě zadání velkého seznamu termínů (týká se spíše bioinformatických aplikací) nebo je-li software právě zaneprázdněn, je možné, že výsledek nebude k dispozici na počkání. Při několikatýdenním testování systému pro účely této publikace se to však stalo pouze v jednom případě.

- PubMatrix vyžaduje registraci, která je však bezplatná a umožňuje např. ukládání hotových matic s možností jejich pozdějšího prohlížení nebo opakovaného prohledávání v časovém odstupu, během kterého mohlo dojít k nárůstu citací ve zdrojové databázi MEDLINE/PubMed. Uvedený limit je tedy spíše výhodou.

3. Závěr

Ohromný nárůst množství vědeckých lékařských informací a současně vysoké požadavky na jejich kvalitu z pohledu medicíny založené na důkazech znesnadňují lékařům cestu k rychlému nacházení správných a prověřených odpovědí na klinické otázky. Tato práce předkládá přehled pěti nástrojů umožňujících efektivní vyhledávání publikací v databázi MEDLINE/PubMed s ohledem na specifika EBM.
Potěšitelná je skutečnost, že existují volně dostupná rozhraní a programy pro efektivní vyhledávání dokumentů relevantních s ohledem na kritéria medicíny založené na důkazech, která pokrývají širokou škálu potřeb a rešeršních schopností uživatelů.

Literatura

[1] G. Guyatt, "Evidence-Based Medicine", *ACP J Club*, vol. 114, p. A-16, 1991.

[2] M. Dawes, W. Summerskill, P. Glasziou et al., "Sicily statement on evidence-based practice", *BMC Medical Education* , vol. 5, p. 1, 2005.

[3] D. L. Sackett, W. M. C. Rosenberg, J. A. M. Gray, et al., "Evidence based medicine: what it is and what it isn't.", *BMJ*, vol. 312, pp. 71–72, 1996.

[4] V. Papíková, "Strategie vyhledávání aktuálních a důkazy podložených informací pro potřeby klinické praxe", *Ikaros*, 2002,
viz `http://ikaros.ff.cuni.cz/2002/c02/med.htm` nebo u autorky
`papikova@euromise.cz`.

[5] V. Papíková, "Medicína založená na důkazech (evidence-based medicine), Cochranova knihovna a rozhodovací procesy v medicíně", *Ikaros*, 2001,
viz `http://ikaros.ff.cuni.cz/2001/c11/med.htm` nebo u autorky
`papikova@euromise.cz`.

[6] D. J. Rouse, "Evidence Based Medicine: underlying Principles", *Clinical Obstetrics and Gynecology*, vol. 41, pp. 235–331, 1998.

[7] W. S. Richardson, M. C. Wilson, J. Nishikawa, R. S. Hayward, "The well-built clinical question: a key to evidence based decisions", *ACP J Club*, vol. 123, pp. A12–13, 1995.

[8] D. L. Sackett et al. "Evidence Based Medicine - How to Practice & Teach EBM", Churchill Livingstone, 1996.

[9] National Institute for Health and Clinical Excellence, "The guidelines manual - Chapter 5: Developing clinical questions", pp. 31–34, 2006,
see `http://www.nice.org.uk/page.aspx?o=308768`.

[10] P. Fontelo, F. Liu, M. Ackerman, "askMEDLINE: a free-text, natural language query tool for MEDLINE/PubMed", *BMC Medical Informatics and Decision Making*, vol. 5, p. 5, 2005,
see `http://www.biomedcentral.com/1472-6947/5/5`.

[11] M. Muin, P. Fontelo, F. Liu, M. Ackerman, "SLIM: an alternative Web interface for MEDLINE/PubMed searches - a preliminary study", *BMC Medical Informatics and Decision Making*, vol. 5, p. 37, 2005,
see `http://www.biomedcentral.com/1472-6947/5/37`.

[12] M. Muin, personal communication, e-mail: mmuin@mail.nih.gov, July 2006.

[13] M. Muin, "PubMed Interact (BETA) Tutorial",
see `http://pmi.nlm.nih.gov/interact/tutorial.html`.

[14] K. G. Becker, D. A. Hosack, G. Dennis Jr, et. al, "PubMatrix: a tool for multiple literature mining", *BMC Bioinformatics*, vol. 4, p. 61, 2003,
see `http://www.biomedcentral.com/1471-2105/4/61`.

Elimination tree based repartitioning

Post-Graduate Student:
ING. TOMÁŠ PĚNIČKA

Department of mathematics
Faculty of Science
MUNI
Janáčkovo náměstí 2a
602 00 Brno

Czech Republic

xpenicka@bart.math.muni.cz

Supervisor:
PROF. ING. MIROSLAV TŮMA, CSC.

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8

Czech Republic

tuma@cs.cas.cz

Field of Study:
Scientific computations

Abstract

This article introduces a novel approach to graph repartitioning. When processing a Cholesky factorisation of sparse SPD matrix, problem of graph partitioning is often involved. In some cases it is needed to repartition the existing components. This operation may be unacceptably costly, because the elimination trees of individual components must be rebuilt. This article describes a novel approach how to process the repartitioning more efficiently.

1. Introduction

A graph partitioning problem is often involved when solving a system of linear equations in parallel environment. This paper focuses on sparse systems with symmetric positive definitive (SPD) matrix. System with SPD matrix can be efficiently solved using the Cholesky factorisation. We suppose that reader is familiar with this method including its sparse modifications in parallel environment.

The paper is organised as follows. Basic terminology is given in section 2. Then some useful assertions about chordal graphs are given in section 3. Then we come to the main point of the paper. Removal of vertex from graph is shown in section 4, insertion of a vertex is shown in 5. A practical implementation of the repartitioning is given in section 6.

2. Basic terminology

In this section we will overview the basic notation used in this paper.

2.1. Required knowledge

First we suppose that reader is familiar with common terminology used in connection with sparse Cholesky factorisation, namely:

- filled matrix, associated (or elimination) graph, filled graph
- elimination order, topological ordering, tree preserving ordering
- chordal graph
- structure of column/vertex
- elimination tree: parent, ancestor, descendant, children, independent vertices, subtree,

- coarse/fine parallelism

If reader need a help with some listed notions [5, 6, 8] will help.

In the whole paper we make the following assumptions:

- The matrix of the problem is sparse and SPD. It is marked A.
- The elimination ordering of A is marked P.
- The filled matrix of A is marked F.
- The associated graph of the matrix A is marked $G(A)$ or simply G.
- The graph of the matrix F (filled graph of A) is marked $G(F)$.
- The elimination tree of A is marked T.
- If G is chordal, the perfect elimination ordering of G (or A) is P.

2.2. Notations and assumptions

To simplify the notation used in the whole paper, we will use the following notation and symbols in the whole paper.

When removing a vertex from G (it is the same as to remove a vertex and column v from A) we suppose in addition to 2.1:

- A vertex which is removed from the graph is marked v. The same letter is used for its position in the elimination ordering.

When inserting a vertex to G (it is the same as to insert a vertex and column v to A) we suppose in addition to 2.1:

- A vertex which is added to the graph is marked v. It is added on the i^{th} position.

It is also needed to identify graph/matrix after removal or addition of a vertex. So we will use the following notation:

- When we insert (i.e. move from an another part of a larger graph) a vertex v into G, we denote the set of its neighbours in G by $\mathrm{Adji}_G(v)$ (ADJacent set In G).
- $G \dashv v$ denotes the graph G with added vertex v and all edges between v and its neighbours from G. (So the sets $\mathrm{Adji}_G(v)$ and $\mathrm{Adj}_{G \dashv v}(v)$ are equal.)
- By $A \dashv (v, i)$ we mean matrix A with row/column corresponding to v inserted on the i^{th} position (the i^{th} and consecutive rows/columns shift one up), the elimination ordering of $A \dashv (v, i)$ is $P \dashv (v, i)$. $G \dashv v$ is its elimination graph.
- By $P \dashv (v, i)$ we mean ordering P extended thus that v is inserted before the i^{th} vertex. After the insertion, i means $i + 1^{st}$ position etc.
- By $F \dashv (v, i)$ ($T \dashv (v, i)$) we mean filled matrix (elimination tree) of $A \dashv (v, i)$.
- By $P \setminus v$, we mean ordering P without vertex v (i.e. P restricted to the vertex set without v).
- $A \setminus v$ means A without line and row corresponding to v.
- $G \setminus v$ ($F \setminus v$, $T \setminus v$, respectively) is the associated graph (elimination tree, filled graph, respectively) of $A \setminus v$.

2.3. Used symbols

We will use he following symbols:

- $\mathrm{PARENT}_T(x)$ means parent of x in T.
- $\mathrm{Struct}_L(x)$ ($\mathrm{Struct}_T(x)$) means structure of a column (vertex) in matrix L (tree T).
- $\mathrm{Adj}_G(x)$ means set of vertices adjacent to x in G.
- $\mathrm{Adji}_G(v)$ means set of vertices adjacent to v in G before the vertex v is inserted into G.
- $\mathrm{Children}_T(x)$ means set of children of x in T.
- $\mathrm{Ancestor}_T(x)$ means a set of ancestor of x in T, including x. If there is a set of vertices instead of single x, it means union over ancestors of the individual elements of the set.
- $T[x]$ means a subtree of T rooted in $x \in T$.

If there cannot arise a confusion we often omit the subscripts.

3. Removal and addition of elements for chordal graphs

When describing the changes connected with repartitioning of the graph, we suppose, that the graph G is a component of a bigger graph. Then after repartitioning it may happen that some vertex is removed from G or is added to G.

It turns out that it is important whether the graph is chordal or not. In case of chordal graphs it is much more simpler to update the elimination tree. So we will care about chordality of the graph after removal or addition of a vertex or an edge.

When removing a vertex from G, we can ask following question: First, is the graph $G \setminus\setminus v$ chordal? And second, is the ordering of A with omitted v a perfect elimination ordering for $A \setminus\setminus v$? The answer to both questions is positive as stated in the following theorems.

Theorem 1 *Let the basic assumptions hold (section 2). Let G be is chordal. Then the following assertions hold.*

1. *$G \setminus\setminus v$ is a chordal graph.*
2. *$A \setminus\setminus v$ is a perfect elimination matrix and $P \setminus\setminus v$ is its perfect elimination ordering.*

Remark 1 *If $G \setminus\setminus v$ is not connected, then theorem 1 also holds with a slight modification of item 1: components of $G \setminus\setminus v$ are chordal.*

When inserting a vertex, it is not so easy to resolve whether the new graph is chordal. The following theorem gives a condition under which the new graph is chordal.

Theorem 2 *Let the basic assumptions hold (section 2). Let G be a chordal graph. Let v be inserted to G. Then to check the chordality of $G +\!\!\!+ v$ it is sufficient to check only cycles which include the vertex v.*

In the following theorem we will state some simpler conditions under which the graph $G +\!\!\!+ v$ is chordal.

Theorem 3 *Let the basic assumptions hold (section 2). Let G be a chordal graph. Let v be inserted into G. Then $G +\!\!\!+ v$ is chordal if any of the following conditions is fulfilled.*

1. $\mathrm{Adji}_G(v)$ *is a clique, i.e.,* v *is simplicial in* $G \dotplus v$.

2. *A subgraph induced by vertices* $\mathrm{Adji}_G(v)$ *is chordal. There exists a clique* $K \subseteq \mathrm{Adji}_G(v)$ *such that for every vertex* w *from* $\mathrm{Adji}_G(v) \setminus K$ *it follows* $\mathrm{Adj}_G(w) \subseteq \mathrm{Adji}_G(v)$.

Unlike in case of removal of vertex, it is not clear whether after insertion of a vertex the new ordering is perfect. It turns out that generally it is not true. It may happen that for any i the new ordering $P \dotplus (v, i)$ is not perfect. Even any topological reordering cannot improve this situation. The following theorem gives a sufficient condition for the new elimination order to be perfect. Compare to theorem 3.

Theorem 4 *Let the basic assumptions hold (section 2). Let* G *be a chordal graph. Let vertex* v *be inserted into* G *on the* i^{th} *position. Then* $P \dotplus (v, i)$ *is perfect if any of the following conditions is fulfilled.*

1. $\mathrm{Adji}_G(v)$ *is a clique and the vertices in the clique are in* P *numbered higher then or equal to* i.

2. *Vertices from* $\mathrm{Adji}_G(v)$ *which are numbered larger than or equal to* i *form a clique (denote it by* K*) and subgraph induced by vertices* $\mathrm{Adji}_G(v)$ *is chordal and the elimination ordering* P *restricted to these vertices is perfect.*

Now let us continue with a very important theorem which gives the necessary condition for $G \dotplus v$ to be chordal.

Theorem 5 *Let the basic assumptions hold (section 2). Let* G *be chordal. The necessary condition for* $G \dotplus v$ *to be chordal is that subgraph of* G *induced by vertices* $\mathrm{Adji}_G(v)$ *is connected.*

4. Removal of a vertex for general graph

Let us explore, what are the problems connected with removing of a vertex from the associated graph of A.

Suppose we have no need to change the elimination ordering together with removal a vertex. We only omit the removed vertex from the given ordering P.

The simplest way how to explore the changes after removal the vertex is to look at the matrix A. If a vertex v is removed, then the corresponding row and column disappear from A, see figure 1.

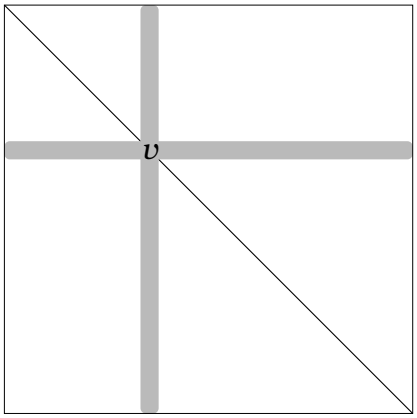


Figure 1: Removal of a vertex. A Column and a row are highlighted by the grey colour disappear from the matrix A.

When looking at the matrix L, removal of the appropriate row and column is not the only change in the matrix, some additional fills may also disappear. Before we analyse possible cases, we will show a useful example. Consider matrix A_1 with the filled matrix F_1, where $\circ$ marks the additional fill

$$F_1 = \begin{pmatrix} v & & \times & & \times & & \times & \times & & \\ & 1 & & & & \times & & & & \\ \times & & 2 & & \circ & & \circ & \times & & \\ & & & 3 & & & \times & & & \\ \times & & \circ & & 4 & & \circ & \circ & \times & \\ & \times & & & & 5 & & & \times & \\ \times & & \circ & \times & \circ & & 6 & \circ & \times & \times \\ \times & & \times & & \circ & & \circ & 7 & \circ & \times \\ & & & & \times & \times & \times & \circ & 8 & \times \\ & & & & & & \times & \times & \times & 9 \end{pmatrix}$$

Then a matrix $\widetilde{F}_1 = F_1 \text{“} v$ looks like

$$\widetilde{F}_1 = \begin{pmatrix} 1 & & & & \times & & & & \\ & 2 & & & & & \times & & \\ & & 3 & & & \times & & & \\ & & & 4 & & & & \times & \\ \times & & & & 5 & & & \times & \\ & & \times & & & 6 & & \times & \times \\ & \times & & & & & 7 & & \times \\ & & & \times & \times & \times & & 8 & \times \\ & & & & & \times & \times & \times & 9 \end{pmatrix}$$

Notice, that $\widetilde{F}_1$ has no additional fill. In figure 2, the elimination tree for both matrices is shown.

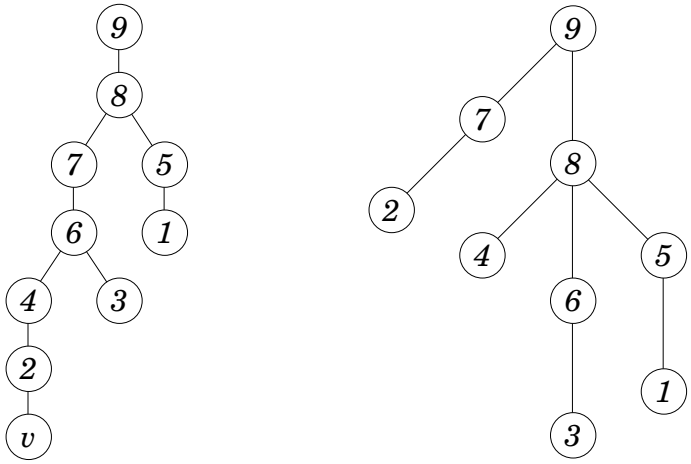


Figure 2: Elimination trees of F_1 (left) and $\widetilde{F}_1$ (right).

What happened in this example? The additional fills from 2 disappeared, so 7 became parent of 2. Then 4 connected to 8 instead of 6. Then 6 lost its additional fill and connected to 8 instead of 7. Finally 7 without additional fill imported originally from v connected directly to 9. The removed leaf of the elimination tree caused changes up to the root.

The following useful lemma tells us how the structure of columns of L is changed.

Lemma 1 *Let the basic assumptions hold (section 2). Then for every column x from $A \text{“} v$ it holds:* $\mathrm{Struct}_{T \text{“} v}(x) \subset \mathrm{Struct}_T(x)$.

So the following corollary can be easily derived:

Corollary 1 *Let the basic assumptions hold (section 2). Let v be removed from G.*
If $y \in \mathrm{Ancestor}_{T \ominus v}(x)$ then $y \in \mathrm{Ancestor}_T(x)$, in particular
if $y = \mathrm{PARENT}_{T \ominus v}(x)$ then $y \in \mathrm{Ancestor}_T(x)$.

The corollary can be used to find where the changes in the elimination tree are located. It is summarised in the following theorem which gives a theoretical ground to the practical application of elimination tree based repartitioning.

Theorem 6 *Let the basic assumptions hold (section 2). Remove a vertex v from T. Then the following assertions hold:*

1. *Structures of vertices which are not in $\mathrm{Ancestor}_T(\mathrm{Adj}_G(v))$ are not modified by the removal of v.*
2. *Two vertices cannot swap the roles in their ancestor–descendant relationship. In particular, if x was an ancestor of y in T, it cannot happen that y is an ancestor of x in $T \ominus v$.*

The following theorems give an upper bound on location on changes in the elimination tree.

Theorem 7 *Let the basic assumptions hold (section 2). Remove a vertex v from G. Let us define set N as:*

$$N = \bigcap_{i \in \mathrm{Adj}_G(v)} \mathrm{Ancestor}_{T \ominus v}(i).$$

Then the structures of the vertices from the set N are unchanged.

This statement is only of theoretical interest. We want to compute an elimination tree $T \ominus v$ by small changes of T, but the estimate of range of the changes demands the knowledge of still unknown tree $T \ominus v$. At least it is possible to state the following observation.

Theorem 8 *Let the basic assumptions hold (section 2). Let $g \in \mathrm{Ancestor}_T(v)$. If $\mathrm{Struct}_T(g) = \mathrm{Struct}_{T \ominus v}(g)$, then there are no changes in the structures of vertices from $\mathrm{Ancestor}_{T \ominus v}(g)$.*

Example 3 *Consider the elimination tree from figure 2. After removal of v, according to theorem 6, vertices, which are not ancestors of $2, 4, 6$ or 7 , i. e. vertices from $T[3]$ and $T[5]$ don't change their structure. Vertex 8 doesn't change its structure. Notice that 7 became independent of 8 and connected directly to 9. It is not in contradiction with theorem 8*

5. Insertion of a vertex

When inserting a vertex we have a freedom of choice of the order of v. An important question is where to insert v into sequence of ordered vertices? We will show, what happens if we insert the vertex v on a particular position in the elimination ordering.

Denote neighbours of v in G by $i_1 < i_2 < \cdots < i_k$ (numbers come from the initial ordering of matrix A).

5.1. Insertion before neighbours

The following theorem gives us an information, where v will be connected to the tree after insertion the column v before its neighbours.

Theorem 9 *Let the basic assumptions hold (section 2). Let $\mathrm{Adj}_{G+v}(v) = \{i_1 < i_2 < \cdots < i_k\}$, v is inserted on the i^{th} position and $i \leq i_1$. Then*

1. *v is a leaf in $T+(v,i)$.*
2. *i_1 is the parent of v in $T+(v,i)$ (formally $\mathrm{PARENT}_{T+(v,i)}(v) = i_1$).*

After connecting of v we would like to know, how the whole tree is changed:

Theorem 10 *Let the basic assumptions hold (section 2). Let $\mathrm{Adji}_G(v) = \{i_1 < i_2 < \cdots < i_k\}$, let v be inserted on the i^{th} position and $i \leq i_1$. Then $\mathrm{Ancestor}_T(\mathrm{Adji}_G(v))$ form one line in $T+(v,i)$. Moreover*

$$\mathrm{Ancestor}_T(\mathrm{Adji}_G(v)) = \mathrm{Ancestor}_{T+(v,i)}(\mathrm{Adji}_G(v)) = \mathrm{Ancestor}_{T+(v,i)}(v)$$

The rule that a child is numbered before its parent still holds.

An interesting situation occurs, when we reorder the vertices with the tree preserving ordering before inserting the vertex v. It may happen that the result differs from the previous one. This experience is summarised in the next observation.

Observation 1 *A reordering which is tree preserving with respect to T, is generally not tree preserving with respect to $T+(v,i)$. (Generally, it does not depend particularly on i.)*

This property can be well exploited when we want to affect the ordering of vertices from $\mathrm{Ancestor}_{T+(v,i)}(u_1, \ldots u_k)$ which were in disjoint subtrees (in T). This is convenient for reducing of fill connected in the new tree.

The following theorems give an upper bound on location on changes in the elimination tree. The first one more theoretical and gives an answer in advance before the computation is performed, while the second can be considered as a dynamic stopping criterion.

Theorem 11 *Let the basic assumptions hold (section 2). Let us define set N as follows*

$$N = \bigcap_{i \in \mathrm{Adji}_G(v)} \mathrm{Ancestor}_T(i),$$

then the structure of vertices from N remains unchanged after addition of v.

Theorem 12 *Let the basic assumptions hold (section 2). Let $g \in \mathrm{Ancestor}_T(v)$. If $\mathrm{Struct}_T(g) = \mathrm{Struct}_{T-v}(g)$, then there are no changes in the structures of vertices from $\mathrm{Ancestor}_{T-v}(g)$.*

Notice that the vertex g generally is not element of N.

5.2. General case

Now we will focus on general case, where vertex v is not inserted before its neighbours. First we solve a question, which vertices become children of v. The answer is given in the next simple theorem, which complements theorem 9.

Theorem 13 *Let the basic assumptions hold (section 2). Let v be inserted into G. Let $\mathrm{Adji}_G(v) = \{i_1 < i_2 < \cdots < i_k\}$, v is inserted on the i^{th} position and $i_l < i < i_{l+1}$. ($1 < l \leq k$). Then the children of v are given by*

$$\bigcup_{j=1}^{l} c_j \mid c_j = \max\{d \mid d \in \mathrm{Ancestor}_G(i_j) \text{ and } d < i\}. \tag{1}$$

Simply speaking if we order ancestors of i_j, with numbers smaller than i, in an increasing order then the last vertex connects to v (in other words it is a child of v).

The vertices from $\mathrm{Ancestor}_T(i_1, \ldots, i_l)$ with numbers larger than i appear in the structure of v, in addition to $i_l + 1, \ldots, i_k$, which appear in the structure. So the situation above v is similar to theorem 10.

Theorem 14 *Let the basic assumptions hold (section 2). Let* $\mathrm{Adji}_G(v) = \{i_1 < i_2 < \cdots < i_k\}$, v *is inserted on the* i^{th} *position and* $i_l < i \leq i_{l+1}$. *(*$1 < l \leq k$*). Following sets of vertices align into one line and compose the (proper) ancestor set of* v*:*

1. $\mathrm{Ancestor}_T(i_1, \ldots i_l)$ *with number higher than or equal to* i*.*
2. $\mathrm{Ancestor}_T(i_{l+1}, \ldots i_k) = \mathrm{Ancestor}_T(\mathrm{Struct}(T + v)v)$

The vertices could be written in one set as an $\mathrm{Ancestor}_T(i_1, \ldots, i_k)$ with number higher than or equal to i. We divided the vertices into two groups for clarity and for comparison with theorem 10.

The next theorem gives an information about the structure of vertices after insertion of a vertex v.

Theorem 15 *Let the basic assumptions hold (section 2). Let us insert* v *into* T *on the* i^{th} *position. Then following statements hold:*

1. *Only vertices from* $\mathrm{Ancestor}_T(\mathrm{Adji}_G(v))$ *may change their structure.*
2. *Vertices from* $\mathrm{Ancestor}_T(\mathrm{Adji}_G(v))$ *numbered lower than* i *are descendants of* v*, the others are its ancestors.*

6. Practical application

In this section we will show algorithm how to recompute the structure of elimination tree after insertion of vertex. It means to recompute parent, children, and structure of each changed vertex. The algorithm for removal of a vertex is similar but there is not enough space in this paper to state both of them.

Suppose we are inserting v on the i^{th} position and we know the adjacent set of v in G ($\mathrm{Adji}_G(v)$), let us denote it by $\mathrm{Adj}(v)$. Define following sets:

$$C = \text{set of children of } v \text{ defined by equation 1}$$

Next define

$$D = \text{vertices from } \mathrm{Ancestor}(\mathrm{Adj}(v)) \text{ with number} < i,$$

they become the subset of the Descendants of v ($C \subseteq D$). Last define

$$E = v \text{ and vertices from } \mathrm{Ancestor}(\mathrm{Adj}(v)) \text{ without } D \text{ and } N,$$

N is from theorem 11, these vertices become the subset of the ancestors of v (see theorem 15).

Example 4 *Let's focus on example in figure 3. There is an elimination tree, where we want to insert* v *to the* 19^{th} *position.* $\mathrm{Adji}_G(v)$ *is marked by grey colour. The elimination order is written by the circles. In the circle, there is written a set (or sets) which the vertex belongs to. Neighbours of* v *with number* $< i$ *are* $\{5, 12\}$, $c_5 = 15,\ c_{12} = 17$. *For* i_{12} *is* $j = 15$ *The result of the insertion is in figure 4.*

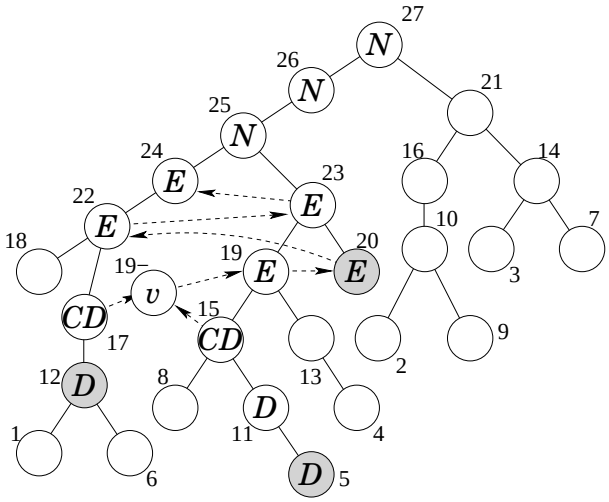


Figure 3: Demonstration of sets C, D, E when inserting a vertex. $\mathrm{Adji}_G(v)$ is marked with grey colour.

Now we can proceed to the algorithm.

Algorithm 1 *Recomputation of elimination tree after inserting of a vertex*

1. *Add* v *to the structure of all vertices in* D.
2. *For all* $c \in C$ *assign* $\mathrm{PARENT}(c) \leftarrow v$.
3. *Sort* E *by numbers in an increasing order, denote them by* $v = e_0, e_1, \ldots e_{q-1}$. *Denote* $\mathrm{PARENT}(e_{q-1})$ *as* e_q.
4. *For all vertices* $e_1, \ldots, e_q$ *remove from their children sets vertices which belong to* $C \cup E$.
5. *Assign* $\mathrm{Children}(e_0) \leftarrow C$.
6. *For* i *from* 1 *to* q *do*
   (a) *Find* $\mathrm{Struct}(e_{i-1})$
   (b) *Assign* $\mathrm{PARENT}(e_{i-1}) \leftarrow e_i$.
   (c) *Add* e_{i-1} *to* $\mathrm{Children}(e_i)$.

7. Conclusion

This paper shows that it is possible to repair elimination tree after addition or removal of vertices/edges from graph without need to rebuild it from scratch. The algorithm how to recompute basic properties of individual vertices was shown. Author will work on integration of the algorithms into some solver of systems of linear equations to compare its performance with traditonal techniques.

References

[1] Timothy A. Davis and William W. Hager *Modifying a Sparse Cholesky Factorization*, SIAM J. Matrix Anal. Appl. Vol. 20, No. 3, pp. 606-627, 1999.

[2] Michele Benzi, José Marín and Miroslav Tůma *A Two-Level Parallel Preconditioner Based on Sparse Approximate Inverses*, in: Iterative Methods in Scientific Computation IV, D. R. Kincaid and A.

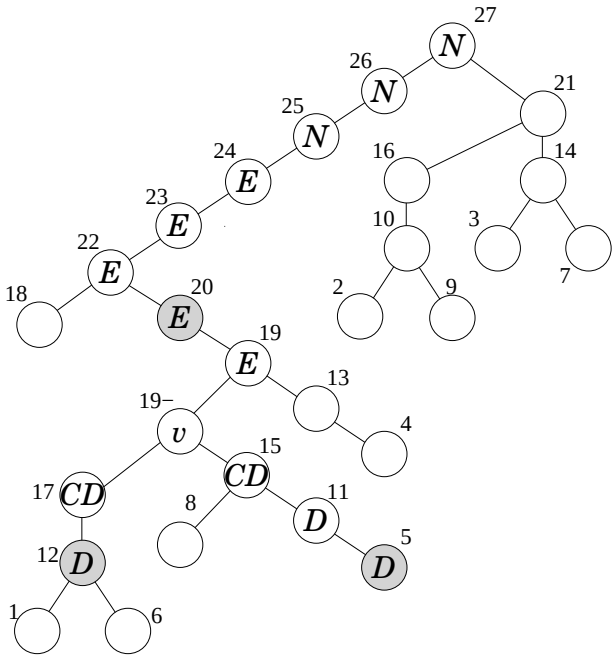


Figure 4: Demonstration of sets C, D, E when removing a vertex. Notice, that N must be learned from figure 3.

C. Elster, eds., IMACS Series in Computational and Applied Mathematics, Vol. 5, IMACS, New Brunswick, NJ (1999), pp. 167–178.

[3] Michele Benzi and Miroslav Tůma *A Comparative Study of Sparse Approximate Inverse Preconditioners,* Applied Numerical Mathematics 30 (1999), pp. 305–340.

[4] Michele Benzi and Miroslav Tůma *Orderings for Factorized Sparse Approximate Inverse Preconditioners,* SIAM J. Sci. Comput. 21 (2000), pp. 1851–1868.

[5] John R. Gilbert, Donald J. Rose and Anders Edenbrandt *A Separator theorem for Chordal Graphs,* SIAM J. Alg. Disc. Meth. 5 (1984), pp. 306–313.

[6] Martin Charles Golumbic *Algorithmic Graph Theory and Perfect Graphs,* Academic Press, New York, 1980.

[7] Joseph W. H. Liu *Reordering Sparse Matrices for Parallel Elimination,* Parallel Computing 11 (1989), pp. 73–91.

[8] Joseph W. H. Liu *The Role of Elimination Trees in Sparse Factorization,* SIAM J. Matrix Anal. Appl. 11 (1990), pp. 134–172.

Some remarks on bidiagonalization and its implementation

Post-Graduate Student:
ING. MARTIN PLEŠINGER

Faculty of Mechatronics
Technical University of Liberec
Hálkova 6
461 17 Liberec 1
Czech Republic

martin.plesinger@tul.cz

Supervisor:
PROF. ING. ZDENĚK STRAKOŠ, DRSC.

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic

strakos@cs.cas.cz

Field of Study:
Computational mathematics

This work was supported by the National Program of Research "Information Society" under project 1ET400300415.

Abstract

In this contribution we are interested in orthogonal bidiagonalization of a general rectangular matrix. We describe several bidiagonalization algorithms and focuse particularly on implementation details.

The bidiagonalization is the first step in hybrid methods, which are usually used to solve ill-posed and rank-deficient problems [5] arising for instance from image deblurring. Thus we consider a matrix from ill-posed problem and illustrate the behavior of discussed bidiagonalization methods on this example.

The second example points on the very close relationship between bidiagonalization and the so called core theory in linear approximation problems [4], see also [10]. The relationship with the Lanczos tridiagonalization of symmetric positive semidefinite matrices analyzed in [7], [8] is mentioned.

1. Introduction

Consider a general rectangular matrix $A \in \mathbb{R}^{n\times m}$. The orthogonal transformations

$$P^T A Q = \begin{bmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \beta_3 & \alpha_3 & \\ & & \ddots & \ddots \end{bmatrix} = B, \quad R^T A S = \begin{bmatrix} \beta_1 & \alpha_1 & & \\ & \beta_2 & \alpha_2 & \\ & & \beta_3 & \alpha_3 \\ & & & \ddots & \ddots \end{bmatrix} = C,$$

where P, Q, R, S have mutually orthonormal columns, are called the **lower** and the **upper bidiagonalization**, respectively. Obviously these transformations are closely connected, e. g., the upper bidiagonalization of A can be expressed as the lower bidiagonalization of A^T.

In the rest of this section only the lower bidiagonalization is considered, the results for the upper bidiagonalization are analogous. Define the **full bidiagonal decomposition**

$$P^T A Q = B \in \mathbb{R}^{n\times m}, \quad \text{where} \quad P \in \mathbb{R}^{n\times n}, \quad Q \in \mathbb{R}^{m\times m} \tag{1}$$

are orthogonal matrices and $P = [\,p_1\,,\,\ldots\,,\,p_n\,]$, $Q = [\,q_1\,,\,\ldots\,,\,q_m\,]$. Consequently $A = P B Q^T$. Depending on the relationship between n and m, the matrix B can contain a zero block. Furthermore, define the k-th **partial bidiagonal reduction**, $k \leq \min\{n, m\}$,

$$P_k^T A Q_k = B_k \in \mathbb{R}^{k\times k}, \quad \text{where} \quad P_k \in \mathbb{R}^{n\times k}, \quad Q_k \in \mathbb{R}^{m\times k}, \tag{2}$$

B_k is the upper left principal submatrix of B and P_k, Q_k contains first k columns of P, Q, respectively. Thus

$$B = \left[\begin{array}{c|cc} B_k & & \\ \hline e_k^T \beta_{k+1} & \alpha_{k+1} & \\ & \beta_{k+2} & \alpha_{k+2} \\ & & \ddots \quad \ddots \end{array}\right], \quad P = [P_k\,,\, p_{k+1}\,,\, \ldots\,,\, p_n], \quad Q = [Q_k\,,\, q_{k+1}\,,\, \ldots\,,\, q_m].$$

Note that in general $A \neq P_k\, B_k\, Q_k^T$. Finally, define the $(k+)$-th partial bidiagonal reduction, $k < \min\{n, m\}$,

$$P_{k+1}^T\, A\, Q_k = \left[\frac{B_k}{e_k^T \beta_{k+1}}\right] = B_{k+} \in \mathbb{R}^{(k+1)\times k}, \quad \text{where} \quad P_{k+1} \in \mathbb{R}^{n\times(k+1)}, \quad Q_k \in \mathbb{R}^{m\times k}.$$

Remark 1 *Similarly, the k-th or $(k+)$-th upper partial bidiagonal reduction produces bidiagonal matrices $C_k \in \mathbb{R}^{k\times k}$, $C_{k+} \in \mathbb{R}^{k\times(k+1)}$, respectively.*

2. Two approaches to bidiagonalization

In this section two basic tools for computation of the bidiagonal transformation are summarized – the Householder method, which leads to the full bidiagonal decomposition, and the Golub-Kahan algorithm, which leads to the partial bidiagonal reduction. At the end of this section the connection between these approaches is explained.

2.1. Householder method

The first technique can also be called the **bidiagonalization using orthogonal transformations**. It is well known, that for any vector x an orthogonal matrix G can be constructed such that $Gx = e_1\|x\|_2$. Using these matrices, the matrix A can be transformed in the following way:

$$\left[\begin{array}{cccc} \bullet & * & * & * \\ \bullet & * & * & * \\ \bullet & * & * & * \\ \bullet & * & * & * \end{array}\right] \to \left[\begin{array}{cccc} \clubsuit & \bullet & \bullet & \bullet \\ & * & * & * \\ & * & * & * \\ & * & * & * \end{array}\right] \to \left[\begin{array}{cccc} \clubsuit & \clubsuit & & \\ & \bullet & * & * \\ & \bullet & * & * \\ & \bullet & * & * \end{array}\right] \to \ldots \to \left[\begin{array}{cccc} \clubsuit & \clubsuit & & \\ & \clubsuit & \clubsuit & \\ & & \clubsuit & \clubsuit \\ & & & \clubsuit \end{array}\right],$$

i. e., A is multiplied alternately from the left and right by matrices $G_1, G_2^T, G_3, G_4^T, \ldots$, where the matrices G_j are block-diagonal with two blocks – the first block is an identity matrix of growing dimension, the second block is the above described matrix G. The vectors x are marked by "$\bullet$" in each step, the length of the vectors is decreasing. Denote $\bar{S}_1 \equiv I_m$,

$$\bar{R}_k \equiv G_1^T\, G_3^T\, G_5^T\, \ldots\, G_{2k-1}^T \in \mathbb{R}^{n\times n}, \quad \bar{S}_{k+1} \equiv G_2\, G_4\, G_6\, \ldots\, G_{2k} \in \mathbb{R}^{m\times m}, \tag{3}$$

the orthogonal matrices. Define the k-th and $(k+)$-th **incomplete bidiagonal decomposition**

$$\bar{R}_k^T\, A\, \bar{S}_k = \left[\begin{array}{c|c} C_{(k-1)+} & \\ \hline e_1\, e_k^T\, \beta_k & A^{(k+)} \end{array}\right] = \bar{C}_k, \quad \bar{R}_k^T\, A\, \bar{S}_{k+1} = \left[\begin{array}{c|c} C_k & e_k\, e_1^T\, \alpha_k \\ \hline & A^{(k+1)} \end{array}\right] = \bar{C}_{k+}, \tag{4}$$

respectively, where $A^{(k+)}$ and $A^{(k+1)}$ are the nonbidiagonalized remains of the original matrix, see Section 3. Consequently $A = \bar{R}_k\, \bar{C}_k\, \bar{S}_k^T$, $A = \bar{R}_k\, \bar{C}_{k+}\, \bar{S}_{k+1}^T$. Matrices G may be the Householder reflection matrices, or composed Givens rotation matrices, see, e. g., [3]. Usage of the Householder reflections is discussed in details in Section 3.

The bidiagonalization using orthogonal transformations is useful if the full bidiagonal decomposition of A is required. On the other hand, if only a partial transformation is computed, which is often the case, this approach can be very ineffective. We refer to this technique as **Householder method** (or decomposition).

2.2. Golub-Kahan algorithm

The second technique is based on the **Golub-Kahan bidiagonalization** algorithm [2] (also called the Lanczos bidiagonalization, or the Golub-Kahan-Lanczos bidiagonalization). This approach is different from the previous one. Assuming that the bidiagonal elements are nonnegative, the Householder decomposition is uniquely determined. The outputs of the Golub-Kahan algorithm, i. e. the bidiagonal form of A, depend on the vector $b \in \mathbb{R}^n$. The algorithm follows:

```
00:  β1 := ‖b‖2 ;        u1 := b/β1 ;
01:  wL := A^T u1 ;
02:  α1 := ‖wL‖2 ;       v1 := wL/α1 ;
03:  wR := A v1 ;
04:  β2 := ‖wR‖2 ;       u2 := wR/β2 ;
05:  for   j = 2, 3, 4, ...
06:        wL := A^T uj − v_{j−1} βj ;
07:        αj := ‖wL‖2 ;      vj := wL/αj ;
08:        wR := A vj − uj αj ;
09:        β_{j+1} := ‖wR‖2 ;  u_{j+1} := wR/β_{j+1} ;
10:  end .
```

The algorithm is obviously stopped on the first $\alpha_j = 0$ or $\beta_j = 0$. Consider that $\alpha_1, \ldots, \alpha_k$ and $\beta_1, \ldots, \beta_{k+1}$ are nonzero (and thus positive). It is easy to show that $\{u_j\}_{j=1}^{k+1} \subset \mathbb{R}^n$, $\{v_j\}_{j=1}^{k} \subset \mathbb{R}^m$ are two sets of mutually orthonormal vectors. Denote

$$P_j \equiv U_j = [\, u_1, \ldots, u_j \,] \in \mathbb{R}^{n \times j}, \quad Q_j \equiv V_j = [\, v_1, \ldots, v_j \,] \in \mathbb{R}^{m \times j}.$$

Then

$$P_k^T A Q_k = \begin{bmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \beta_k & \alpha_k \end{bmatrix} = B_k, \quad P_{k+1}^T A Q_k = \left[\begin{array}{c} B_k \\ \hline e_k^T \beta_{k+1} \end{array}\right] = B_{k+}.$$

The Golub-Kahan algorithm returns in k-th iteration, at the line `07` and `09`, the k-th and $(k+)$-th lower partial bidiagonal reduction of A, respectively.

This algorithm is useful if the partial bidiagonal reduction is required. We refer to this algorithm as the **Golub-Kahan algorithm**.

Remark 2 *The Golub-Kahan bidiagonalization of the matrix A starting from the vector b is very closely related to the Lanczos tridiagonalization of the matrices AA^T, and A^TA with starting vectors $b/\|b\|_2$ and $A^Tb/\|A^Tb\|_2$ respectively, see, e. g., [1]; and to the Lanczos tridiagonalization of the augmented matrix*

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}, \quad \textit{with the starting vector} \quad \begin{bmatrix} b/\|b\|_2 \\ 0 \end{bmatrix}.$$

For more information about this relationship and its connection to the so called core problem, see [7], [8].

2.3. Connection between Householder and Golub-Kahan technique

Consider the k-th partial lower bidiagonalization $B_k = P_k^T A Q_k$ of A computed by the Golub-Kahan algorithm starting from the vector b. Remember that $\beta_1 = \|b\|_2$. It can be easily proven that the matrix

$$P_k^T \left[\begin{array}{c|c} b & A \end{array}\right] \left[\begin{array}{c|c} 1 & \\ \hline & Q_k \end{array}\right] = \left[\begin{array}{c|ccc} \beta_1 & \alpha_1 & & \\ & \beta_2 & \alpha_2 & \\ & & \ddots & \ddots \\ & & & \beta_k \quad \alpha_k \end{array}\right] = \left[\begin{array}{c|c} e_1\,\beta_1 & B_k \end{array}\right] \tag{5}$$

is the left upper principal submatrix of $(k+)$-th incomplete upper bidiagonal decomposition and thus the submatrix of matrix $\bar{C}_{k+}$, returned by the Householder method applied on the extended matrix $[\,b\,|\,A\,]$. Here $\bar{R}_k = [\,P_k\,,\,\bar{r}_{k+1}\,,\,\ldots\,,\,\bar{r}_n\,]$, and $\bar{S}_{k+1} = [\,[\,\mathrm{blockdiag}\,(1, Q_k)\,]\,,\,\bar{s}_{k+2}\,,\,\ldots\,,\,\bar{s}_{m+1}\,]$.

3. Implementation

In the previous section, two theoretical approaches to computation of the bidiagonalization were discussed. In this section, some implementation details of bidiagonalization algorithms are analyzed. We briefly describe these procedures, their advantages and disadvantages. The emphasis is put on influence of rounding errors, computational cost, and also computation speed and memory requirements.

3.1. Householder method [HH]

The implementation of the Householder method corresponds to the procedure described above. First, the Householder matrices are defined, then the algorithm follows.

Definition 1 *The Householder (orthogonal) matrix, which transforms the vector $x_1 \in \mathbb{R}^n$ into the vector $x_2 \in \mathbb{R}^n$, $\|x_1\|_2 = \|x_2\|_2 \neq 0$, $x_1 \neq x_2$, has the form*

$$H(x) = I - 2\,\frac{x x^T}{\|x\|_2^2} \in \mathbb{R}^{n\times n}, \quad \textit{where} \quad x = x_1 - x_2\,. \tag{6}$$

Thus $H(x)\,x_1 = x_2$, $H(x)\,x_2 = x_1$ and $H(x) = H(-x) = (H(x))^T$.

Denote $A^{(1)} \equiv A$. Denote a_1 the first column the matrix $A^{(j)}$ and $\tilde{a}_1$ the first row of the matrix $A^{(j+)}$ (matrices $A^{(j)}$ and $A^{(j+)}$ are generated during the computation). The algorithm follows:

00: generate a zero matrix $\bar{C} := 0 \in \mathbb{R}^{n\times m}$;
01: generate two identity matrices $\bar{R} := I_n$ and $\bar{S} := I_m$;
02: **for** $j = 1, 2, 3, \ldots$
03: $\beta_j := \|a_1\|_2$; $\bar{c}_{j,j} := \beta_j$;
04: $H_{2j-1} := H(\,a_1 - e_1\,\beta_j\,) \in \mathbb{R}^{(n-j+1)\times(n-j+1)}$;
05: remove a_1 from $A^{(j)}$; $A^{(j+)} := H_{2j-1}\,A^{(j)} \in \mathbb{R}^{(n-j+1)\times(m-j)}$; *// update of $A^{(\cdot)}$*
06: $\bar{R} := \bar{R}\,[\,\mathrm{blockdiag}\,(\,I_{j-1}\,,\,H_{2j-1}\,)\,]$; *// update of $\bar{R}$*
07: $\alpha_j := \|\tilde{a}_1\|_2$; $\bar{c}_{j,j+1} := \alpha_j$;
08: $H_{2j} := H(\,\tilde{a}_1^T - e_1\,\alpha_j\,) \in \mathbb{R}^{(m-j)\times(m-j)}$;
09: remove $\tilde{a}_1$ from $A^{(j+)}$; $A^{(j+1)} := A^{(j+)}\,H_{2j} \in \mathbb{R}^{(n-j)\times(m-j)}$; *// update of $A^{(\cdot)}$*
10: $\bar{S} := \bar{S}\,[\,\mathrm{blockdiag}\,(\,I_j\,,\,H_{2j}\,)\,]$; *// update of $\bar{S}$*
11: **end**.

The dimensions of matrices $A^{(\cdot)}$ are decreasing during the computation. The algorithm can be stopped for any $j = k$ at the line `06` or `10`, and the small matrix $A^{(k+)}$ or $A^{(k+1)}$, respectively, is written into the right bottom corner of the matrix $\bar{C}$ (this is the nonbidiagonalized part of the original matrix A, see (4)). Then the algorithm yields the k-th or $(k+)$-th incomplete upper bidiagonal decomposition, respectively.

The updates in the lines `05`, `06`, `09`, `10` are not realized by a matrix-matrix products (e. g., $A := A\,H$) which are computationally very expensive, see remark 3. If the knowledge of the structure (6) of the Householder matrix H is used, then the multiplication can be rewritten in the form

$$A\,H \;=\; A\left(\,I \;-\; 2\,\frac{x x^T}{\|x\|_2^2}\,\right) \;=\; A \;-\; \frac{2}{\|x\|_2^2}\,(Ax)\,x^T\,. \tag{7}$$

For detailed description, see [3, paragraphs 5.1.3 and 5.1.4].

Remark 3 *The update realized by matrix-matrix product has computational cost of order $\mathcal{O}(n^3)$ the update (7) has computational cost only of order $\mathcal{O}(n^2)$. The implementation with matrix-matrix multiplication is inapplicable, see Figure 1. Further, update by matrix-matrix product is not as stable as (7) update.*

Advantages. Stable implementation (the most stable implementation mentioned here); very quick implementation; orthogonality in the matrices $\bar{R}$, $\bar{S}$ is perfect; in each step the incomplete decomposition is obtained.

Disadvantages. Dense square matrices $\bar{R}$, $\bar{S}$ must be stored.

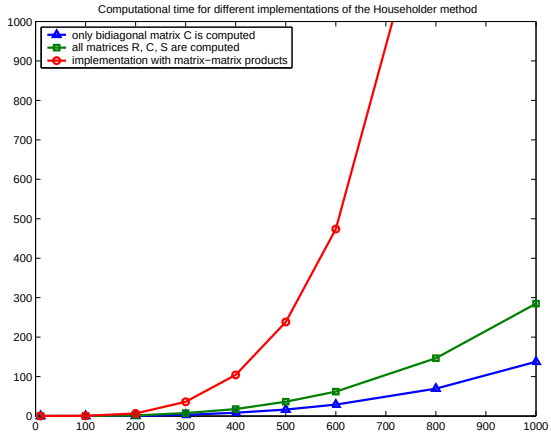


Figure 1: The horizontal axis shows the dimension of a random square matrix, the vertical axis shows the computational times for different implementations of the Householder method. Note that for the implementation with matrix-matrix products and $n = 1000$ the computational time is $t = 3337.9$ seconds.

3.2. Implementation details of Householder method

(*i*) If only the matrix $\bar{C}$ (or its bidiagonal part) has to be computed, the orthogonal matrices $\bar{R}$, $\bar{S}$ need not to be stored (and updated). Consequently, the lines `06` and `10` are leaved out and the algorithm significantly accelerates, see Figure 1.

(*ii*) If the matrix A is considerably rectangular, i. e. $n \gg m$ or $n \ll m$, it is useful to compute at first the QR or LQ decomposition of A, respectively. Assume for example that $n \gg m$ and we want to transform A into the upper bidiagonal form. First, the QR factorization

$$A \;=\; \Pi\,U\,, \quad \Pi \,\in\, \mathbb{R}^{n\times m}\,, \quad U \,\in\, \mathbb{R}^{m\times m}\,,$$

where Π has orthonormal columns and U is upper triangular is computed, then the incomplete (or full) bidiagonal decomposition $U \;=\; \bar{R}\,\bar{C}\,\bar{S}^T$ is found. Summarizing, the corresponding bidiagonal form of A

is the following

$$\bar{\bar{R}}^T A \bar{S}^T = \bar{C} \in \mathbb{R}^{m\times m}, \quad \text{where} \quad \bar{\bar{R}} = \Pi \bar{R} \in \mathbb{R}^{n\times m}, \quad \bar{S} \in \mathbb{R}^{m\times m},$$

and $A = \bar{\bar{R}} \bar{C} \bar{S}^T$. Thus this bidiagonalization is a decomposition.

This QR data preprocessing reduces the computational cost of the bidiagonalization algorithm. The matrix $\bar{\bar{R}}$ is not square and thus we have to be careful while manipulating with it. Obviously, the memory requirements are lower than in the previous algorithms. The QR factorization can be computed by the modified Gram-Schmidt algorithm with iterative refinement. Then the stability of whole procedure is held. For more information, see [3, paragraph 5.4.4].

3.3. Golub-Kahan algorithm without reorthogonalization [GK]

In the rest of this section, we study several implementations of the Golub-Kahan algorithm. All following implementations produce the partial bidiagonal reduction. In general, these algorithms do not get the decomposition of A, because they decompose only the projection of A on a Krylov subspace of growing dimension. The first simple implementation corresponds exactly to the algorithm presented in Section 2.2.

Advantages. The fastest implementation mentioned here; memory requirements are smaller than for the Householder method (if a small part of bidiagonal matrix is computed, we work only with several columns of the matrices P and Q); very simple implementation.

Disadvantages. Total lost of orthogonality in the columns of P and Q ; the bidiagonal elements are computed inexactly, this implementation is in fact inapplicable.

3.4. Golub-Kahan algorithm with reorthogonalization [GK_R]

Better results for the Golub-Kahan bidiagonalization is obtained if the currently computed vectors w_L (line `06`) and w_R (line `08`) are reorthogonalized against previous vectors v_{j-1} , v_{j-2} , $\ldots$, or u_j , u_{j-1} , $\ldots$, respectively.

Really stable implementation is obtained only if the vectors w_L and w_R are reorthogonalized against all previous vectors two times, see [9]. One time reorthogonalization or reorthogonalization against only i previous vectors is not sufficient. More than two times reorthogonalization is not necessary. It has been observed that two full reorthogonalizations is sufficient to achieve results comparable with the Householder method. We refer to this algorithm [GK_R].

Advantages. The orthogonality and stability is as good as in the Householder approach; computational time is comparable with the Householder approach; the memory requirements can be significantly smaller than in the Householder approach (especially if a small part of bidiagonal matrix is computed).

Disadvantages. Computation cost grows with iteration number.

4. Bidiagonalization of ill-posed problem

Consider the matrix

$$A = \text{SHAW}\,(100) \in \mathbb{R}^{100\times 100},$$

and the corresponding right-hand-side vector $b \in \mathbb{R}^n$, from Regularization Toolbox [6]. The system $Ax = b$ is ill-posed, singular values of A decay fast to zero [5] and A is strongly ill-conditioned. Although this matrix is nonsingular, its numerical rank (number of singular values greater than $100\,\|A\|_2\,\epsilon_M$) is equal to 20. We try to bidiagonalize this matrix using the above discussed procedures. All methods are implemented such that they return the same bidiagonal elements in exact arithmetic, i. e. the Householder procedures compute upper bidiagonalization of the extended matrix $[\,b\,|\,A\,]$, the Golub-Kahan procedures starts from the vector b and compute lower bidiagonalization, see Section 2.3.

Figure 2 shows the elements of the vector

$$[\,\beta_1 \,=\, \|b\|_2\,,\,\alpha_1\,,\,\beta_2\,,\,\alpha_2\,,\,\ldots\,,\,\beta_{100}\,,\,\alpha_{100}\,]\,.$$

obtained by different bidiagonalization algorithms. It shows the large difference between elements computed by [HH] and [GK]. Further, we can see four notable peaks (places of total lost of orthogonality between generated vectors) in data computed by [GK] with one time reorthogonalization against only $i = 20$ previous vectors, but only one notable peak if $i = 50$. This total lost of orthogonality disappears if [GK] with full reorthogonalization is used. It is obvious, that [GK] without full double reorthogonalization is useless here. The results of [GK_R] algorithm (with full double reorthogonalization) are not plotted here

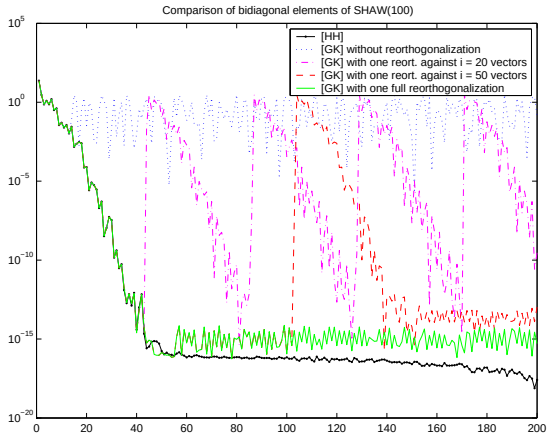


Figure 2: Bidiagonal elements of matrix SHAW(100) computed by different bidiagonalization algorithms.

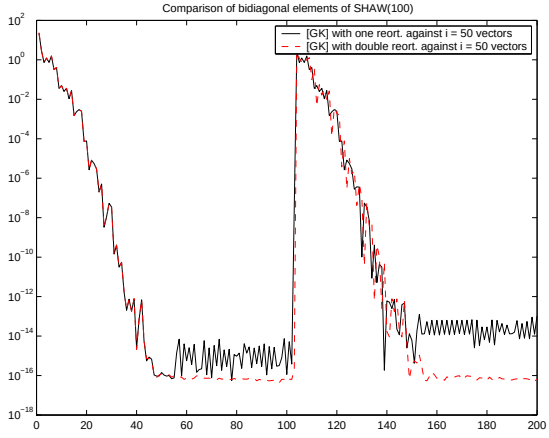


Figure 3: The double reorthogonalization effect used in [GK_R] procedure.

because they are very close to [HH]. In fact, the difference between [HH] and [GK_R] is approximately 5.9494×10^{-13} (note that the difference between [HH] and [GK] with five full reorthogonalizations is approximately 5.4101×10^{-13}). For better illustration of the influence of double reorthogonalization, see Figure 3, where the results of [GK] algorithm with one and two reorthogonalizations against $i = 50$ previous vectors is presented.

The matrix SHAW has very poor numerical properties and thus the differences between the presented algorithms are so large. For matrices having better numerical properties, [GK] with one full reorthogonalization and sometimes also with non-full reorthogonalization can be used. Then the differences between computed bidiagonal forms may not be so significant.

5. The core problem revealing test

Consider an orthogonally invariant linear approximation problem $Ax \approx b$. In [4] it was proved that the upper bidiagonalization of the matrix $[b, A]$ determines the so called core approximation problem that contains all necessary and sufficient information for solving the original problem. If the bidiagonalization is stopped at the first $\beta_{q+1} = 0$ or $\alpha_{q+1} = 0$, then it determines the core problem $B_q\, y = \beta_1\, e_1$ (compatible) or $B_{q+}\, y \approx \beta_1\, e_1$ (incompatible), respectively, see (5). In this section, we consider linear systems which contain the compatible core problem of known dimension q and thus we expect that $\beta_{q+1} = 0$. All the following testing problems are solved by [HH].

5.1. Testing problem A

Consider the matrix and the right-hand-side

$$A = \Pi_1 \left[\begin{array}{c|c} \mathrm{diag}\,(\sigma_1, \ldots, \sigma_q) & 0 \\ \hline 0 & \texttt{rand(n-q)} \end{array}\right] \Pi_2^T \in \mathbb{R}^{n\times n}, \quad b = \Pi_1 \left[\begin{array}{c} \texttt{rand(q,1)} \\ \hline 0 \end{array}\right] \in \mathbb{R}^n,$$

where Π_1, Π_2 are "random" orthogonal matrices computed by `qr(rand(n))` command (see MATLAB manual [11] for description of `qr()`, `rand()`).

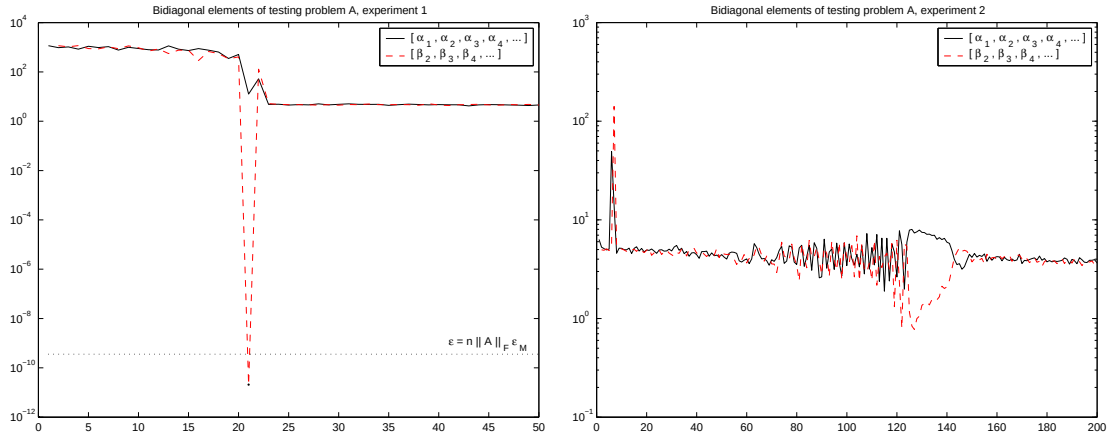


Figure 4: Testing problem A, experiment 1 (*left*): Core problem of dimension $q = 20$ is revealed, $\beta_{21} \approx 10^{-11}$. Experiment 2 (*right*): Core problem of dimension $q = 100$ is not revealed.

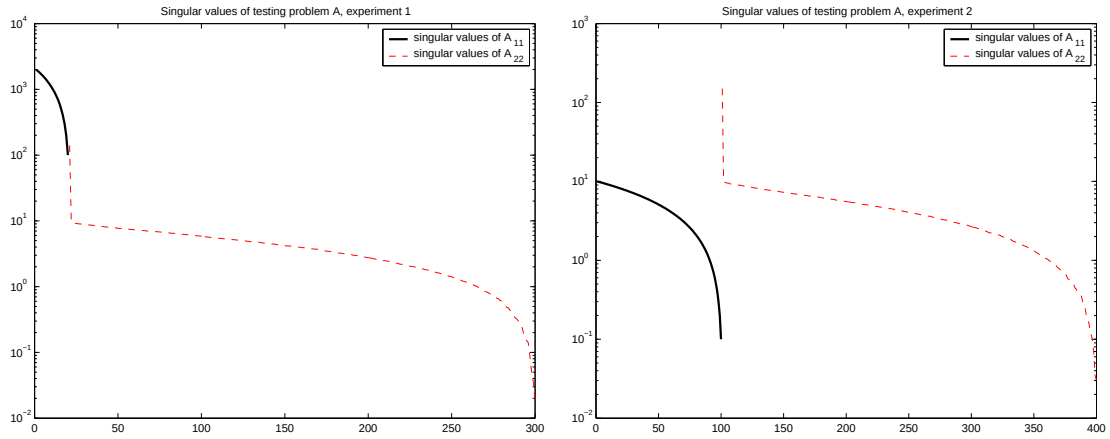


Figure 5: Testing problem A, experiments 1 (*left*) and 2 (*right*): Singular values of $A_{11} = \mathrm{diag}\,(\sigma_1, \ldots, \sigma_q)$ followed by singular values of $A_{22} =$ `rand(n-q)`.

In the first experiment $n = 300$, $q = 20$, $(\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_{20}) = (2000, 1900, 1800, \ldots, 100)$. Bidiagonal elements obtained for this example are plotted in the left graph in Figure 4. In the second experiment $n = 400$, $q = 100$, $(\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_{100}) = (10, 9.9, 9.8, \ldots, 0.1)$. Obtained results are plotted in the right graph in Figure 4. In the first case, the core problem reveals, β_{21} is evidently smaller than the others bidiagonal elements. In the second case, the results are not satisfactory.

In fact, the wrong results obtained for Experiment 2 are the consequence of an inappropriate choice of testing problem. The random matrix $A_{22} =$ `rand(n-q)` has one very large singular value, approximately $\sigma_{\max}(A_{22}) \approx 0.5\,(n-q)$ and the remaining singular values decay linearly, see Figure 5. If $\sigma_{\max}(A_{22}) \gg \sigma_1$, by influence of rounding errors, the dominant singular value of A_{22} is quickly absorbed into the bidiagonal submatrix. This is the consequence of large sensitivity of this problem, despite the stability of computation is held (this will be obvious from the following testing problem).

Remark 4 *We hope, that it will be possible to explain this effect through the connection between the Golub-Kahan bidiagonalization and the Lanczos tridiagonalization, using the properties of Jacobi matrices, see [7], [8].*

5.2. Testing problem B

Consider the matrix and the right-hand-side

$$A = \Pi_1 \begin{bmatrix} \alpha_1 & & \\ \beta_2 & \alpha_2 & \\ & \ddots & \ddots \end{bmatrix} \Pi_2^T \in \mathbb{R}^{n\times m}, \quad b = \Pi_1 \begin{bmatrix} \beta_1 = \texttt{20*rand} \\ 0 \\ \vdots \end{bmatrix} \in \mathbb{R}^n .$$

For $m < n$, sequences $\{\alpha_j\}_{j=1}^{m}$, $\{\beta_j\}_{j=2}^{m+1}$ are generated by the command

```
-10*sort(-rand(m,1)) + rand(m,1);
```

and similarly for $m \geq n$ (see MATLAB manual [11] for description of `sort()`). Both these sequences contain random numbers from the interval $(0, 10)$ approximately sorted from the largest to the smallest number (all numbers are positive). Matrices Π_1, Π_2 are generated in the same way as in the testing problem A. Then we put $\beta_{q+1} = 0$. Summarizing, the n by m rectangular system with compatible core problem of dimension q is created. Moreover, the bidiagonal form of the core problem is known beforehand.

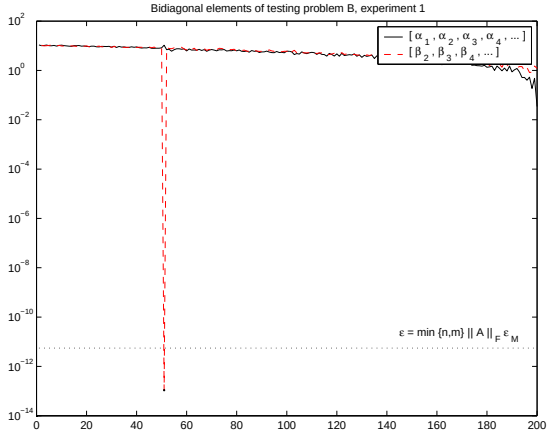


Figure 6: Testing problem B, experiment 1: Core problem is revealed, $\beta_{51} \approx 10^{-13}$.

In the first experiment $n = 1000$, $m = 200$, $q = 50$, in the second experiment $n = m = 1000$, $q = 50$. The core problem reveals without any complication in both cases, for results obtained in the first experiment see Figure 6. Because the bidiagonal form of these problems was predetermined, the absolute

error of the whole process (orthogonal transformation by Π_1 , Π_2 and the following bidiagonalization) can be computed. Denote $B_j \in \mathbb{R}^{j\times j}$ the upper left principal part of the original bidiagonal matrix and $\tilde{B}_j$ the same block of the computed matrix, then

$$\|B_q - \tilde{B}_q\|_2 \doteq 8.704253 \times 10^{-14}, \quad \text{in the first experiment and}$$
$$\|B_q - \tilde{B}_q\|_2 \doteq 5.908292 \times 10^{-14}, \quad \text{in the second experiment,} \quad q = 50\,.$$

The bigger absolute error in the first (smaller) problem is probably caused by the inner QR factorization.

Remark 5 *The absolute error must be computed only from the core problem part of bidiagonal matrix. In finite arithmetic $\tilde{\beta}_{q+1} \neq 0$, but it is very close to zero. The computation does not stop and data obtained during the following computation are strongly influenced by errors.*

6. Calling procedures in MATLAB

All presented procedures were implemented in MATLAB such that they return (theoretically) the same bidiagonal form.

6.1. Householder-based procedures

The [HH] algorithm is called by the commands

```
[P,B,Q] = bidiag_hholder(A,b,k,p,tol,qrf);
      B = bidiag_hholder(A,b,k,p,tol,qrf);
```

where A is a rectangular matrix, b is a vector (e. g., right-hand-side of the problem $Ax = b$). The procedure computes the upper incomplete decomposition of $[\,b \,|\, A\,]$ (and returns the lower decomposition of A). Here k is the number of iterations and p equals 0 or 1. If $p = 0$, then the algorithm produces a square bidiagonal matrix of dimension $k \times k$, if $p = 1$, then it produces a rectangular matrix of dimension $(k + 1) \times k$.

Default values for k and p are $k = \min\{n, m\}$, $p = 0$ if $n \leq m$, or $p = 1$ if $n > m$. With this default values the full decomposition is computed. If only k is prescribed then the default value is $p = 0$ and the square bidiagonal form is returned.

The procedure notifies on each bidiagonal element smaller than `tol`, default value is $2^{-52} \doteq 2.22 \times 10^{-16}$. By the parameter `qrf` we can forbid or enforce the QR factorization of $[\,b \,|\, A\,]$, or LQ factorization of A , default (and recommended) is automatic detection. The decision is based on the dimensions of A , and on the demand to compute only the bidiagonal matrix B , or all matrices P , B , Q .

6.2. Golub-Kahan-based procedures

The Golub-Kahan-based methods [GK], [GK_R] are called by the command

```
[P,B,Q] = bidiag_lgk(A,b,k,p,tol,i,t);
```

This procedure computes a lower partial bidiagonal reduction of A starting from the vector b . The parameters k, p and `tol` and their default values are same as in `bidiag_hholder()`. This procedure returns also the k-th or $(k+)$-th lower partial bidiagonal reduction of A where $p = 0$ or $p = 1$, respectively.

The parameter i specifies the number of vectors against which we want to reorthogonalize, the parameter t specifies how many times the reorthogonalization is performed. The choice

$i = 0$	or	$t = 0$	corresponds [GK] without reorthogonalization,
$i > 0$	and	$t = 1$	corresponds [GK] with reort. against i previous vectors,
$i = -1$	and	$t = 1$	corresponds [GK] with full reorthogonalization,
$i = -1$	and	$t = 2$	corresponds [GK_R] = [GK] with full double reort. (default values).

Another combinations are possible too, but they are not so interesting, e. g., choice

$i = 50$	and	$t = 2$	corresponds the method presented in Figure 3,
$i = -1$	and	$t = 5$	corresponds the method with five times full reorthogonalization.

All numerical examples were carried out on computer Hewlett-Packard Compaq nx9110, Intel Pentium 4 CPU, 2.80 GHz, 448 MB RAM, in MATLAB 6.5 release 13.

References

[1] Å. Björck, "A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations", *BIT*, 28, pp. 659–670, 1988.

[2] G. H. Golub, W. Kahan, "Calculating the singular values and pseudo-inverse of a matrix", *SIAM J. Numer. Anal.*, Ser. B 2, pp. 205–224, 1965.

[3] G. H. Golub, C. F. Van Loan, "Matrix Computations", 3rd edition, *The John Hopkins University Press, Baltimore*, 1996.

[4] C. C. Paige, Z. Strakoš, "Core problems in linear algebraic systems", *SIAM J. Matrix Anal. Appl.*, 27, pp. 861–875, 2006.

[5] P. C. Hansen, "Rank-Deficient and Discrete Ill-Posed Problems", *SIAM, Philadelphia*, 1998.

[6] P. C. Hansen, "Regularization Tools – version 3.2 for MALTAB 6.0", a package for analysis and solution of discrete ill-posed problems, (`http://www2.imm.dtu.dk/~pch/Regutools/index.html`).

[7] I. Hnětynková, Z. Strakoš, "Lanczos tridiagonalization and core problems", to appear in *Lin. Alg Appl.*

[8] I. Hnětynková, M. Plešinger, Z. Strakoš, "Lanczos tridiagonalization, Golub-Kahan bidiagonalization and core problem", to appear in *PAMM*, vol. 6.

[9] G. Meurant, Z. Strakoš, "The Lanczos and conjugate gradients algorithms in finite precision arithmetic", Acta Numerica, pp. 1–70, 2006.

[10] M. Plešinger, Z. Strakoš, "Core problém v úlohách nejmenších čtverců", *In Proceedings of the Doktorandský den ÚI AV ČR*, pp. 102–108, 2005, (`http://www.cs.cas.cz/hakl/doktorandsky-den/files/2005/dk05proc.pdf`).

[11] The MathWorks, Inc. "MATLAB: Documentation", (`http://www.mathworks.com/access/helpdesk/help`).

Multi-Resolution Visualisation of Image Sets with Self-Organizing Maps

Post-Graduate Student:
ING. PHILIP PRENTIS

Department of mathematics
Faculty of nuclear science and physical engineering
ČVUT
Trojanova 13
120 00 Praha 2

Czech Republic

prentisp@km1.fjfi.cvut.cz

Supervisor:
RNDR. LADISLAV ANDREJ, CSC.

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8

Czech Republic

andre@cs.cas.cz

Field of Study:
Mathematical Engineering

Abstract

This paper discusses the well-known cluster analysis and visualisation tool, the self-organizing map (SOM), and its application in image browsing. For this purpose, colour histogram feature vectors are proposed. The pros and cons of different network sizes are discussed, in particular how they are suited to the purposes of direct browsing and also cluster analysis with U-matrices. The tree-structured SOM [6, 7] is proposed as a method of acquiring multi-resolution mappings of a given input space.

1. Introduction

Self-organizing maps (SOM) are a type of artificial neural network that uses unsupervised learning. They were developed by Kohonen [3] in the eighties and have since been employed successfully in a number of applications; in particular they have been used for cluster analysis and visualising high-dimensional data. One example of such an application is the WEBSOM project [3, 4], which uses SOM to map, organize and browse large document files in a two-dimensional content-addressable space (see figure 1). Another is PicSOM [8, 9], which uses multiple SOM to facilitate content-based image retrieval. Both of these projects allow direct browsing of data described by multi-dimensional feature vectors, which is the chief topic of this paper.

In the author's previous work [1] SOM were used to browse bitmap images sorted according to colour content in mappings of varying dimensionality. This paper shows some of the more interesting results of this study, as well as newer results using high-resolution tree-structured self-organizing maps (TS-SOM) [6, 7], which are used extensively in the above-mentioned PicSOM project.

2. Self-Organizing maps

Simply put, a SOM reduces a high-dimensional input-space (set of vectors from R^N) to a two-dimensional ordered grid of codebook vectors (neurons) $\{n_{ij}|i = 1..N_1, j = 1..N_2; \text{typically} N_1 = N_2\}$ that quantify it. In a successfully adapted map adjacent grid nodes will quantify similar data, i.e. data points that have a small Euclidean distance. The principle may be seen in figure 2, which shows a typical SOM in a 2D input space. A simple description of the basic algorithm may be found in [5].

The algorithm

The self-organization process is achieved as follows:

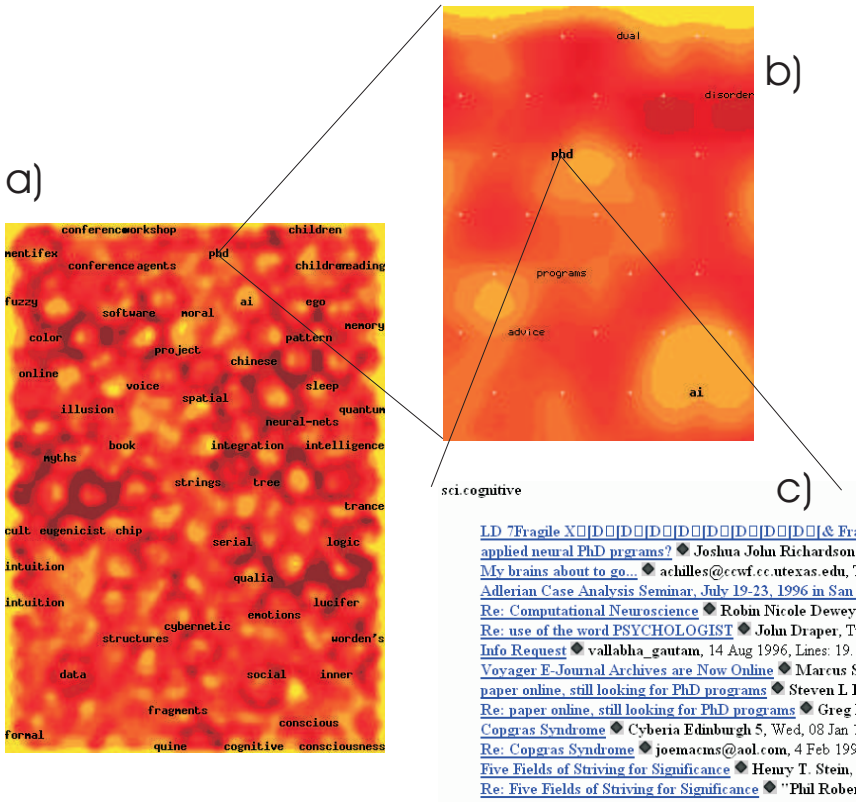


Figure 1: a) A WEBSOM visualization of a large file of documents. The colour represents the file density, yellow/light being a higher density and red/dark being a lower one. Automatically generated labels show topics discussed in documents in different parts of the map. b) A closer view of part of the map. c) A list of documents located at a specific node. These pictures were taken from an online demonstration of WEBSOM at [15].

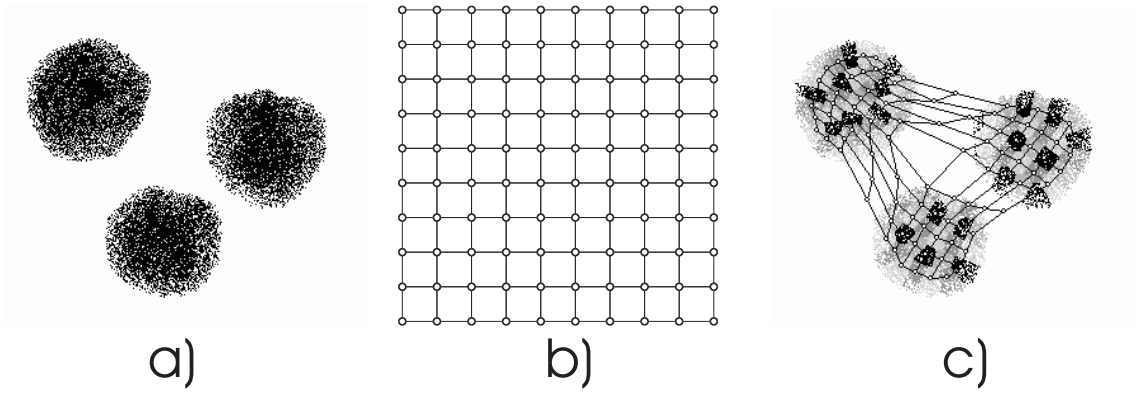


Figure 2: a) 2-dimensional input vectors in 3 clusters. b) A SOM, topologically ordered in a grid of 10×10 neurons c) The SOM adapts itself to match the input space, each neuron's codebook vector quantifying a set of inputs.

1. Initialise the codebook vectors $n_{ij}(0)$ at random (usually by setting them to randomly chosen input vectors).
2. Select a random input $i(t)$ and find the best matching neuron (BMN) $n_{best}(t)$ (i.e. the neuron with the closest codebook vector). Every input sample has the same probability of being selected.
3. Move the BMN and its topological neighbours within a certain neighbourhood distance towards the selected input vector. Units located topologically further from BMN are moved less.

$$n_{ij}(t+1) = n_{ij}(t) + \eta(t) \cdot \phi(i,j,t) \cdot [i(t) - n(t)], \tag{1}$$

where

$$\eta(t) : \ N_0 \rightarrow < 0; 1 > \qquad \text{monotonously decreasing,} \tag{2}$$
$$\phi(i,j,t) : \ N_0 \times N_0 \times N_0 \rightarrow < 0; 1 >,$$

ϕ decreases monotonously with the topological distance of n_{ij} from n_{best} and with t. The topological distance is the length of the shortest path from one neuron to the other in the graph (grid) that represents the network's topology.

4. Proceed to iteration $t+1$. Repeat 2 and 3 iteratively, reducing the proportion of the distance moved η and the neighbourhood distance ϕ each iteration, until they reach a certain predetermined threshold.

As a result the codebook vectors will be attracted to large clusters of input vectors as these will have a higher probability of being selected than sparsely populated areas of input space. η and ϕ must be selected with care if the algorithm is to achieve good results [5]. See [1] for an in-depth heuristic analysis and optimization of these functions.

3. Input space

It is necessary to pre-process input space so that the individual data items are described by feature vectors. Typically, when visualising an N-dimensional input space with a SOM (an $m \times m$ grid), we are faced with the task of displaying m^2 N-dimensional codebook vectors as 2D graphical images in such a way that each image describes the type of data characterised by the given codebook vector. Alternately, we may use the image to describe the input data that has been allocated to the codebook vector through vector quantization. This may be done by finding the mean vector of all the allocated data and working with it rather than the codebook vector. Further possibilities might include using the convex envelope of the data or otherwise describing data distribution at the area of the codebook vector. However, this paper deals with the other simpler alternative, that of describing the actual codebook vectors.

Graphics and colour features

To use SOM visualisation effectively, it is first necessary to convert the image database to feature vectors. If we are using RGB coded (Red, Green, Blue) colour pictures then this conversion F may be defined as follows:

$$F : A \rightarrow R^N, \tag{3}$$

where

$$A \text{ is a matrix } m \times n, A \equiv (a_{ij} \in RGB)_{i=1\ldots m, j=1\ldots n} \tag{4}$$
$$RGB \equiv [0 \ldots 255]^3$$

F should be such a transformation that the resulting feature vector describes how a human perceives the picture. For example, if we take a holiday photo, then a human might describe it as a sandy beach with deep blue water and sky. While the information that the picture contains *a beach*, *water* and *sky* may be difficult to obtain automatically without the use of meta-data, the dominant colours and their respective quantities can be easily obtained from the RGB coding of the picture.

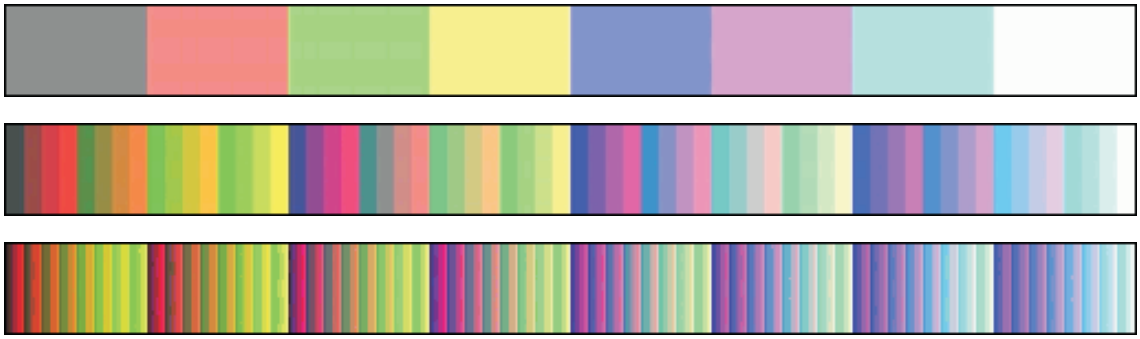


Figure 3: Colour palettes of 8, 64 and 512 colours. These are all cubes of 2, 4 and 8 respectively, which guarantees they will divide colour space (256^3 colours) into bins of equal size.

One simple yet effective method of feature classification [1] is the use of colour histograms. The RGB colour space is segmented into equal-sized bins. Each element of an image's feature vector counts the number of pixels in the picture of a given colour bin. Simply put, the vector says how much of each colour is contained in the picture.

Depending on how many bins we segment colour space into, we can make the feature vector more or less discriminative. Too many bins differentiate images of similar appearance, while too few group dissimilar images together. Heuristically, we have found that 64 colour bins achieve the best results, as they create a palette of colour sets, the mean colours of which are different enough to be registered as separate colours, while the colour variance tolerated in a given set may be considered to be varying shades of the given mean. More bins cause similar colours to be classed differently, while fewer bins produce so little discriminability as to forgo usefulness. See fig. 3.

4. Network resolution

When using a SOM to visualise a data set, it is important to choose a network of the correct size. This may be best described by the neuron/input (NI) ratio.

$$NI = N/I \tag{5}$$

where N is the number of neurons or nodes in the network and I is the number of input vectors being visualised. In this section we will show the differences between using a high or a low input ratio and discuss which are more suitable for different tasks.

4.1. Low NI ratio

For the purposes of this paper $NI < 1$ is considered low. A low NI ratio assures that most or all neurons will quantify at least some input vectors. This will hold truer the lower the ratio and in the extreme case of $NI = 1/I$ will be guaranteed. On average each neuron will quantify $1/NI$ input vectors. If used for data browsing, then we select such a ratio that $1/NI$ is the number of data elements we would typically like to browse at a given moment. If NI is too low, the user may be confronted with too many elements at once causing him to miss the one he was looking for. If it is too high then searching may be difficult, as many of the nodes will quantify only a few elements. Nodes that quantify no elements at all are not useful because they waste the user's time and may be confusing to navigate through.

The goal of the optimisation procedure is to achieve as even a distribution as possible while preserving topology and minimising the average quantisation error (AQE). AQE is the average distance of an element in input space from its BMN. The more we demand a rigid topology, the harder it is to achieve a low AQE. Usually, good topology preserving tends towards a more even distribution [1].

If used correctly, this method allows us to view data in chunks organized by similarity into a map. It is most

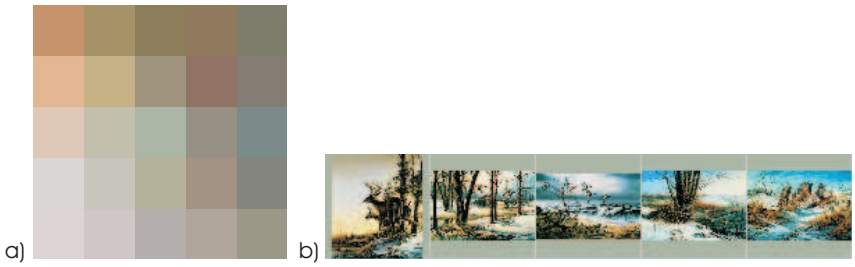


Figure 4: a) A 5×5 SOM visualising a set of bitmap images characterised by colour-histogram vectors. Each square is coloured in the mean colour of the codebook vector's histogram. b) Thumbnails of the images quantified by a selected node, in this case the centre node.

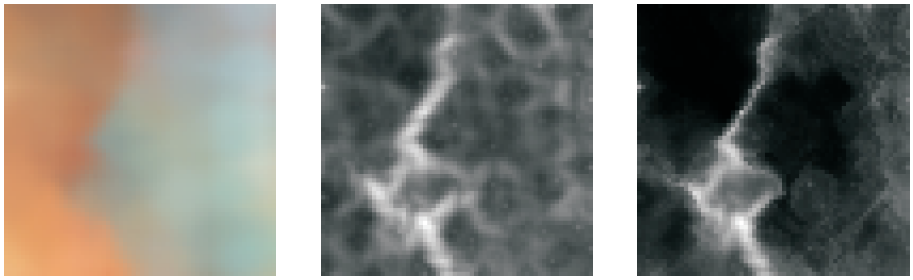


Figure 5: The colour map, U-matrix and U*-matrix of a 64×64 SOM mapping of 121 data elements in approximately 4 clusters. In each of these clusters further sub-clusters may be found. The U-matrix is made up of the average distances of the neurons from their immediate neighbours, black signifying very close and white very distant. Neurons within a cluster will be close to each other while at the edges they will be further apart. The U*-matrix is a U-matrix that has been processed to emphasize the clusters.

suitable for direct browsing. On the downside, the low network resolution does not show clusters very well and is unsuitable for analysing input space as a whole. Figure 4 shows a typical mapping.

4.2. High NI ratio

As discussed in [10], large-scale networks with many neurons may be used to bring out the emergent qualities of the SOM. We observe the overall structure of the SOM rather than the data quantified by individual neurons. One method is to calculate the average distance of each unit from its immediate topological neighbours, creating what is called a U-Matrix.

$$U \equiv [u_{ij}], \tag{6}$$

$$u_{ij} = \frac{1}{\mid X_{ij} \mid} \sum_{k,l \in X_{ij}} \mid n_{ij} - n_{kl} \mid,$$

where

$$X_{ij} \equiv \Big\{ k, l \mid k \in \{1 \ldots N_1\}, l \in \{1 \ldots N_2\}; \mid k - i \mid + \mid l - j \mid = 1 \Big\}$$

If we display a U-matrix graphically as in figure 5, we can see how clusters are being mapped by noting which areas of the matrix have low values. Graphically, these will appear as 'valleys' of nodes inside clusters divided by 'ridges' of borderline nodes. For more information on U-matrices (and their improved variant, the U*-matrix) see [11].

It should be noted that with high NI ratio networks most neurons will not quantify any inputs. Thus, such networks are unsuitable for direct browsing purposes, as a user would have a difficult time finding inputs sparsely distributed across the map.

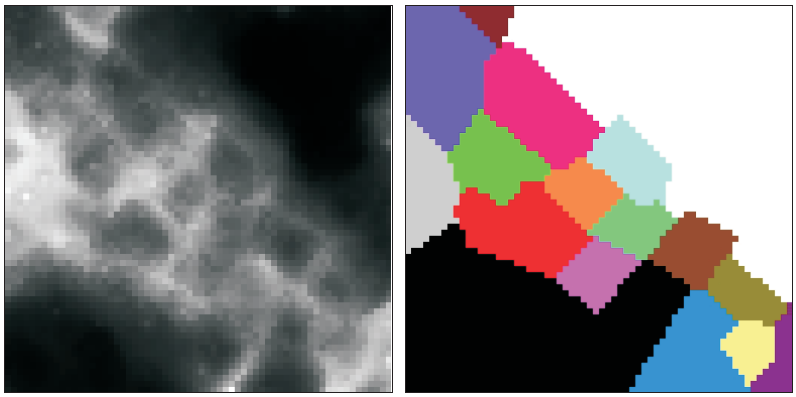


Figure 6: A U*-matrix and the resulting cluster map. These clusters were selected by hand according to the borders visible on the U*-matrix. Ideally, an automatic process would be used to determine the clusters.

Figure 7: A gallery of pictures displayed in a quantum tree-map with Photomesa image browser. Each partition represents a folder. Here, the algorithm is optimised to use available space as efficiently as possible, while enforcing rectangular partitioning.

When using a U- or U*-matrix to locate clusters, we take sets of closely-packed nodes (low U-matrix values) bordered by nodes that are far from their immediate neighbours (high U-matrix values). See figure 6. These clusters, however, are unordered and do not form a regular map.

It would be possible to supply the user with information regarding adjacent clusters or alternatively, the clusters could be displayed using a quantum tree-map with zoomable navigation [12]. Figure 7 shows a large set of images displayed in a quantum tree-map by *Photomesa* image browser.

Unfortunately, larger networks take longer to adapt, as more neurons must be searched when finding the best-matching neuron (BMN). This may be an issue if the SOM is part of a user-application where adaptation speed is paramount.

5. Tree-structured SOM

One way we can simultaneously analyse input space with high and low input ratios is to use a tree-structured self-organizing map (TS-SOM), [6, 7]. This is a hierarchical structure of SOMs of exponentially increasing size. Each level of the TS-SOM adapts separately, but in the lower levels, the search for the best-matching neuron is limited to those hierarchically connected to the BMN of the previous layer. See figure 8.

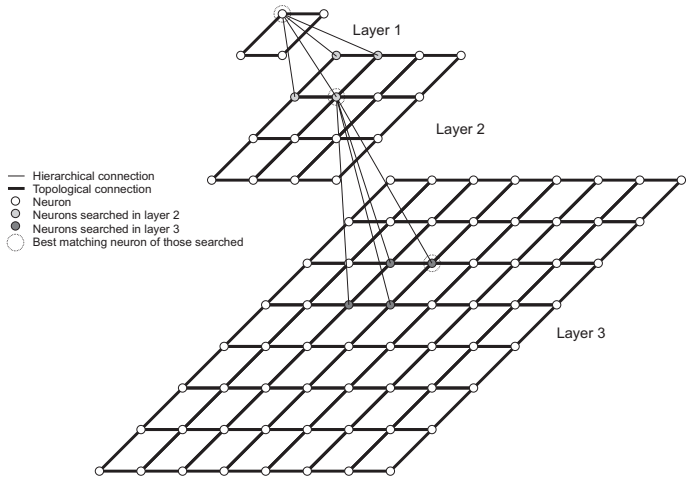


Figure 8: A 3-layer TS-SOM with 4 neurons at top layer and 64 at the bottom

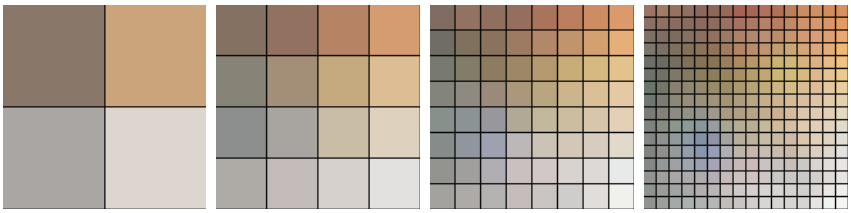


Figure 9: 4 layers of a TS-SOM show colour distribution of an input space of images at different resolutions

5.1. The basic algorithm

The algorithm works as follows:

1. Perform one iteration of the SOM algorithm on the top layer.
2. Perform one iteration of the SOM algorithm on the next layer, but limit the search for the BMU to the neurons located under the winning neuron of the previous layer.
3. Repeat 2 until all layers have been updated.
4. Repeat 1 to 3 until the SOM thresholds have been met.

The advantages of such a structure are obvious. Instead of performing a full-search for the BMN at the lower layers, we restrict ourselves to a constant number of neurons per a given layer, thus greatly increasing the adaptation speed. The complexity of the algorithm is $O(logN)$, where N is the number of neurons on the bottom layer [7]. Also, due to the hierarchical structuring, all the SOMs will be orientated similarly in input space and the TS-SOM as a whole may be considered a multi-resolution mapping of the given data set. See figure 9.

5.2. Wide-search TS-SOM

One unfortunate property of the TS-SOM is the propagation of errors to the lower layers. Inputs bordering between two neurons on a higher layer may gradually become more and more poorly quantified as the BMN becomes more and more restricted. During our experiments we found that on the lower layers, inputs had a tendency to group together rather than spread themselves evenly over the map as is typical in correctly tuned SOMs. As noted in [9], better results may be achieved by allowing searching for the BMN in a wider scope, which includes neurons adjacent to those directly under a higher layer (figure 10).

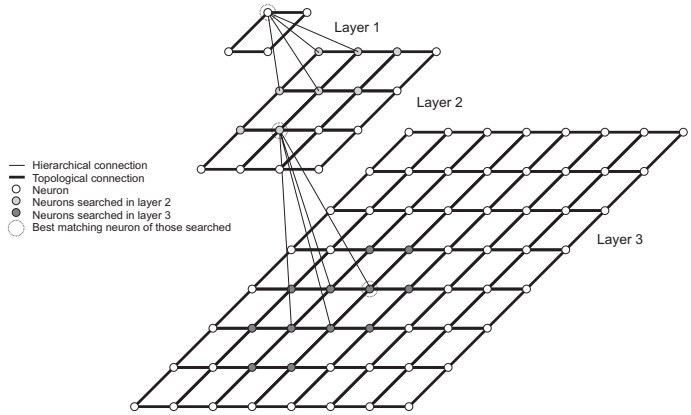


Figure 10: wide-search for the BMN in neighbouring neurons

5.3. Multi-resolution correction

The unfortunate side effect of the wide-search improvement of the TS-SOM algorithm is that the separate layers become desynchronised. This degrades the quality of the TS-SOM as a multi-resolution mapping. In [2] we describe a simple and effective method for rectifying this problem and removing the desynchronisation effect. It should be noted that good synchronisation also improves the quality of the search for the BMN and brings the TS-SOM closer to the quality of the standard full-search methods.

5.4. Alternative models

The TS-SOM is not the only hierarchical model of the SOM. In [13] a *Multi-layer SOM* is proposed. It differs from the TS-SOM in that the higher levels adapt to the positions of the codebook vectors of the level directly beneath them. This means that there is no synchronisation between layers. Also this method does not have the benefit of the TS-SOM's acceleration when seeking the BMN.

Another model is the *Evolving Tree*, which has an indefinite structure that evolves and grows to fit the presented data. It has good data-fitting qualities and retains the speed of the TS-SOM, but due to the unpredictable nature of its structure is unsuitable for direct-browsing applications as described in this paper. See [14].

6. Conclusions and discussion

Self-organizing maps can be used effectively for the purpose of visualising large galleries of colour bitmap images (in [1, 2] we tested mappings of several thousand images). The image pre-processing need not be overly complex, provided that the resulting feature vectors describe the images in a way that a human would perceive them. It is important to take into consideration the neuron/input ratio when designing SOM applications. Small maps are suitable for direct browsing applications, whereas large ones are better for analysing the overall data structure.

By using tree-structured SOM we can achieve both goals at once by creating a synchronized multi-resolution mapping that includes both high and low resolution mappings. However, the lower high-resolution layers will be even less suitable for direct-browsing purposes than standard single-layer SOM of equal dimensions. TS-SOM are, however, advantageous because they greatly accelerate the search for the BMN in large-scale SOM.

Using SOM gives us an advantage over more conventional means of browsing in that they allow us to sort data according to similarity, which facilitates associative searching and may be useful when analysing it. The methods and algorithms described in this paper have been shown mainly in the context of image-browsing, but they could be used just as well for any other kind of data that can be described in terms of

feature vectors.

Future work will consist of developing and exploring new original models and model variants based on self-organizing maps for the purposes of data browsing and analysis.

References

[1] P. Prentis, *Associative Sorting of Data and Kohonen Neural Networks* (original title - *Asociativní třídění dat a kohonenovské neuronové sítě*), Master's thesis. Faculty of Nuclear Science and Physical Engineering, Czech Technical University in Prague. June 2004.

[2] P. Prentis "Multi-Resolution Visualisation of Data with Self-Organizing Maps", being prepared for publication.

[3] T. Kohonen, *Self-Organizing Maps*, Third Edition, Springer-Verlag Berlin 2001.

[4] T. Kohonen, S. Kaski, K. Lagus, J. Salojärvi, J. Honkela, V. Paatero, and A. Saarela, "Self Organization of a Massive Document Collection", *IEEE Transactions on Neural Networks*, Special Issue on Neural Networks for Data Mining and Knowledge Discovery, volume 11, number 3, pages 574-585. May 2000

[5] R. Rojas, *Neural Networks - A Systematic Introduction*, Springer-Verlag Berlin 1996.

[6] P. Koikkalainen, E. Oja, "Self-organizing hierarchical feature maps", Proceedings of *International Joint Conference on Neural Networks*, San Diego, CA, 1990; II:279-284.

[7] P. Koikkalainen, "Progress with the tree-structured self-organizing map", *11th European Conference on Artificial Intelligence*, August 1994.

[8] J. Laaksonen, M. Koskela and E. Oja, "Application of Self-Organizing Maps in Content Based Image Retrieval", in Proceedings of *ICANN'99*, Edinburgh, 1999.

[9] J. Laaksonen, M. Koskela, S. Laakso and E. Oja, "Self-Organizing Maps as a Relevance Feedback Technique in Content Based Image Retrieval", *Pattern analysis & Applications*, 4(2-3): 140-152, June 2001.

[10] A. Ultsch, "Data Mining and Knowledge Discovery with Emergent Self-Organizing Feature Maps for Multivariate Time Series", In E. Oja and S. Kaski, editors, *Kohonen Maps*, pages 33–45, Amsterdam, 1999. Elsevier.

[11] A. Ultsch, "Clustering with SOM: U*C", in Proceedings of the *5th Workshop on Self-Organzing Maps*, Paris 2005.

[12] B. Bederson "PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps", *UIST 2001, ACM Symposium on User Interface Software and Technology, CHI Letters*, 3(2), pp. 71-80. 2001.

[13] J. Koh, M. Suk and S. M. Bhandarkar, "A Multilayer Self-Organizing Feature Map for Range Image Segmentation", *Neural Networks*, 8:67-86, 1995.

[14] J. Pakkanen, J. Iivarinen and E. Oja, "The Evolving Tree - A Novel Self-Organizing Network for Data Analysis", *Neural Processing Letters 20*, 2004.

[15] The WEBSOM research group, "WEBSOM - Self-Organizing Maps for Internet Exploration" `http://websom.hut.fi/websom/`.

Thickness-shear vibration of in-plane parallel piezoelectric resonator

Post-Graduate Student:
ING. PETR RÁLEK

Department of Modelling of Processes
Faculty of Mechatronics and Interdisciplinary Engineering Studies
Technical University of Liberec
Hálkova 6
161 17 Liberec 1

Czech Republic

petr.ralek@tul.cz

Supervisor:
PROF. ING. JIŘÍ MARYŠKA, CSC.

Department of Modelling of Processes
Faculty of Mechatronics and Interdisciplinary Engineering Studies
Technical University of Liberec
Hálkova 6
161 17 Liberec 1

Czech Republic

jiri.maryska@tul.cz

Field of Study:
Science Engineering

Abstract

In the paper, we introduce the application of the mathematical model of piezoelectric resonator. It solves the real problem of thickness-shear vibration of in-plane parallel quartz resonator. The results, compared with the measurement, will be given in the article.

The finite element (FEM) model of the piezoelectric materials, based on linear piezoelectric state equations, leads to a large generalized eigenvalue problem. Resonance frequencies are subsequently found by solving this algebraic problem. Typically, we are not interested in all eigenvalues (resonant frequencies), only for those corresponding to so-called dominant resonant frequencies. For solving the algebraic problem, we use implicitly restarted Arnoldi method implemented in Arpack library. For the coarser meshes, we compute complete spectra and find the frequencies of dominant oscillation modes (according to the electromechanical coupling coefficient). Then we focus on the part of spectra near to the chosen dominant frequency and repeat the computation for refined meshes. From the results, we can also find out the intervals between dominant resonance frequencies (which is other important parameter describing the behavior of resonator).

1. Physical description

We very briefly sketch the physical properties of the piezoelectric materials. The detailed description was given in [2] and [3].

The piezoelectric oscillation oscillation is governed Newton's law of motion and the quasi static approximation of Maxwell's equation. These two equation are accomplished by two linear piezoelectric state equations. The partial differential equation for displacement $\mathbf{u}$ and electric potential φ result,

$$\varrho \frac{\partial^2 \tilde{u}_i}{\partial t^2} = \frac{\partial}{\partial x_j}\left(c_{ijkl}\ \frac{1}{2}\left[\frac{\partial \tilde{u}_k}{\partial x_l} + \frac{\partial \tilde{u}_l}{\partial x_k}\right] + d_{kij}\ \frac{\partial \tilde{\varphi}}{\partial x_k}\right) \quad i = 1,2,3, \tag{1}$$

$$0 = \frac{\partial}{\partial x_k}\left(d_{kij}\ \frac{1}{2}\left[\frac{\partial \tilde{u}_i}{\partial x_j} + \frac{\partial \tilde{u}_j}{\partial x_i}\right] - \varepsilon_{kj}\ \frac{\partial \tilde{\varphi}}{\partial x_j}\right). \tag{2}$$

Quantities c_{ijkl}, d_{kij} and ε_{ij} represent symmetric material tensors, playing role of the material constants. Additional, tensors c_{ijkl} and ε_{ij} are positive definite.

Initial conditions, Dirichlet boundary conditions and Neumann boundary conditions are added:

$$\begin{aligned} \tilde{u}_i(.,0) &= u_i, \quad x \in \Omega, \\ \tilde{u}_i &= 0, \qquad i = 1,2,3, \quad x \in \Gamma_u, \\ \mathrm{T_{ij}} n_j &= f_i, \qquad i = 1,2,3, \quad x \in \Gamma_f, \\ \tilde{\varphi}(.,0) &= \varphi, \\ \tilde{\varphi} &= \varphi_D, \quad x \in \Gamma_\varphi \\ \mathrm{D_k} n_k &= q, \quad x \in \Gamma_q, \end{aligned} \tag{3}$$

where

$$\Gamma_u \cup \Gamma_f = \Gamma,\ \Gamma_u \cap \Gamma_f = \emptyset,\ \Gamma_\varphi \cup \Gamma_q = \Gamma,\ \Gamma_\varphi \cap \Gamma_q = \emptyset.$$

Right-hand side f_i represents mechanical excitation by external mechanical forces, q denotes electrical excitation by imposing surface charge (in the case of free oscillations, they are both zero). Equations (1)-(2) define the problem of harmonic oscillation of the piezoelectric continuum under given conditions (3).

2. Numerical solution

2.1. Weak formulation and discretization

We derive the weak formulation of the problem (1)–(3)in the standard way, using the Green formula and boundary conditions. Then we discretize the problem in space variables, using tetrahedron finite elements with linear base functions. The process of weak formulation and discretization is explained in more details also in [2]. The system of ordinary differential equations for values of displacement and potential in the nodes of division results. It has block structure,

$$\mathrm{M}\ddot{\mathrm{U}} + \mathrm{KU} + \mathrm{P}^{\mathrm{T}}\Phi = \mathrm{F}, \tag{4}$$

$$\mathrm{PU} - \mathrm{E}\Phi = \mathrm{Q}. \tag{5}$$

After introduction of Dirichlet boundary conditions, sub-matrices $\mathrm{M},\ \mathrm{K}$ and E are symmetric and positive definite.

2.2. Generalized eigenvalue problem

The core of the behavior of the oscillating piezoelectric continuum lies in its free oscillation. Free oscillations (and computed eigenfrequencies) tell, when the system under external excitation can get to the resonance. For the free harmonic oscillation, the system (4) can be transformed to

$$\begin{pmatrix} \mathrm{K} - \omega^2\mathrm{M} & \mathrm{P}^{\mathrm{T}} \\ \mathrm{P} & -\mathrm{E} \end{pmatrix} \begin{pmatrix} \mathrm{U} \\ \Phi \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{6}$$

where ω is the circular frequency of oscillation. Eigenfrequencies can be computed by solving the generalized eigenvalue problem

$$\mathrm{AX} = \lambda \mathrm{BX} \tag{7}$$

with

$$\mathrm{A} = \begin{pmatrix} \mathrm{K} & \mathrm{P}^{\mathrm{T}} \\ \mathrm{P} & -\mathrm{E} \end{pmatrix},\ \mathrm{B} = \begin{pmatrix} \mathrm{M} & 0 \\ 0 & 0 \end{pmatrix},\ \lambda = \omega^2,$$

where A is symmetric and B is symmetric and positive semi-definite matrix. Resonance frequency f can be computed from the relation

$$\omega = 2\pi f.$$

Computed eigenvectors (resp. part U) describe the modes of oscillations. For solving the generalized eigenvalue problem (7), we use implicitly restarted Arnoldi method implemented in Arpack library [4] (in Fortran language). Inner steps in the process use direct solver from Skypack library [6] for solving the symmetric indefinite linear systems. It is suitable for solving partial eigenvalue problem (computing of several eigenvalues with high precision) with possibility of the shift and it allows to deal with the sparsity of the matrices. Using of the shift enables to obtain the eigenvalues from desired part of the spectrum with better accuracy.

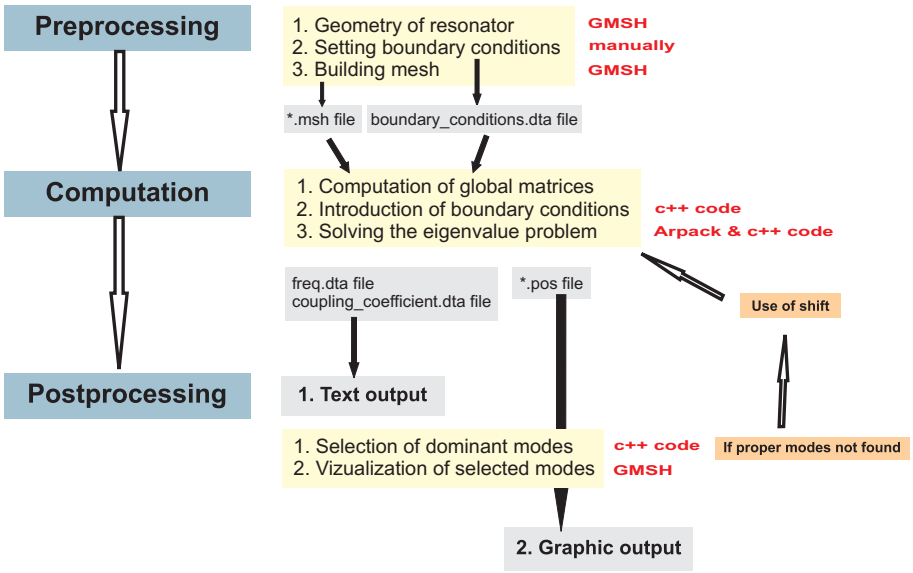


Figure 1: Scheme of computer implementation of the model and its stages.

2.3. Computer implementation

The realization of the model consists of three parts: preprocessing, processing and postprocessing (Fig. 1). In pre- and postprocessing parts, we use the free software GMSH [5] for mesh generation and visualization of the results. For building the global matrices, we developed our own code written in C++ language. For solving the generalized eigenvalue problem, we use the Arpack [4] implementation of implicitly shifted Arnoldi method (in Fortran code). All parts are debugged under Windovs XP.

The preprocessing part consists of building the geometry (according to the engineering assignments) and mesh of the resonator, see Fig. 2. We use the GMSH [5] code. The processing part computes the global matrices and the consecutive eigenvalue problem (using text file with parameters - accuracy, number of computed eigenvalues, shift, etc...). It gives several output files, which are used in the postprocessing. Computed eigenvalues and eigenvectors define the oscillation modes, which are sorted according their *electromechanical coupling coefficients*. The electromechanical coupling coefficient k is defined by the relation [1]

$$k^2 = \frac{E_m^2}{E_{st} E_d},$$

where

$$E_m = \frac{1}{2}\left(\mathrm{U}^T \mathrm{P} \Phi\right), \quad E_{st} = \frac{1}{2}\left(\mathrm{U}^T \mathrm{K} \mathrm{U}\right), \quad E_d = \frac{1}{2}\left(\Phi^T \mathrm{E} \Phi\right)$$

are (sequentially): mutual energy, elastic energy and dielectric energy. The higher the value of k, the better the possibility of excitation of the oscillation mode.

3. Application on the real problem - vibration of in-plane parallel quartz resonator

The model was applied on the thickness-shear oscillation of the in-plane parallel quartz resonator. This resonator is manufactured in the company Krystaly, a.s., resident in Hradec Králové [7]. We looked for dominant modes in x direction.

The geometry and the illustrative mesh of the resonator are shown on the Fig. 2. It is a circular wafer with radius R including two circular electrodes with radius r placed in the middle of upper and bottom side. The resonator is fixed in two opposite spots in x direction. The thickness of the resonator is h. All parameters are listed in the Tab. 1. The resonator is made from quartz $AT - cut_{35.25^\circ}$.

R (mm)	r (mm)	h (mm)
7	4.5	0.333

Table 1: Proportions of resonator sample.

Resonance frequency	Measured (kHz)	Computed (kHz)	Relative difference
basic	4962	4210.843	15.1%
1. harmonic	5067.5 = 1.02 f_1	4246.481 = 1.01 f_1	16.2%
2. harmonic	5102.5 = 1.03 f_1	4331.652 = 1.03 f_1	14.7%

Table 2: Comparison of measured and computed dominant resonant frequencies of thickness-shear vibrational modes.

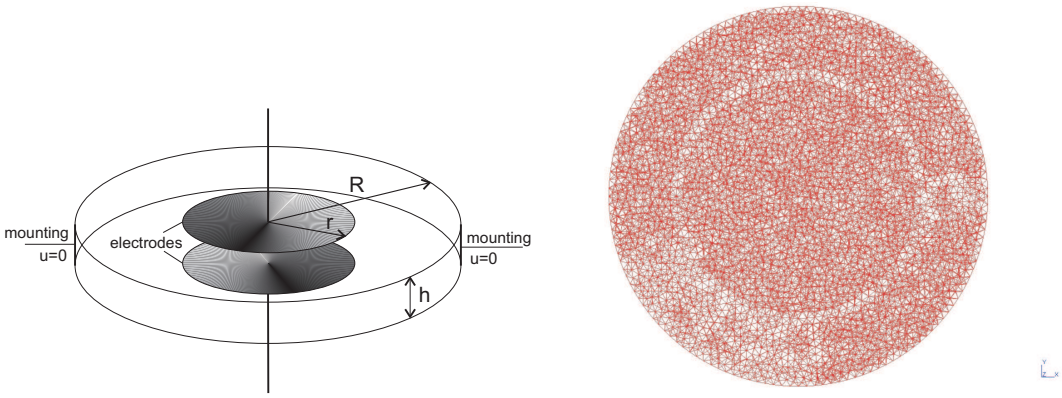


Figure 2: Geometry and illustrative mesh of plan parallel resonator.

For listed results, parameters of the computation were[1]:

Mesh: number of nodes = 7360, number of elements =31860
Matrices:
matrix $\mathbb{A}$ $9771 \times 9771,$ 100430 nonzero elements of upper tringular part,
matrix $\mathbb{B}$ 47038 nonzero elements of upper tringular part
Computational time:
preprocessing part $\sim 3\ min,$ processing part (for 400 eigenvalues) $\sim 55\ min,$
postprocessing part $\sim 1\ min$

Memory demands (for 400 eigenvalues) ~ 0.8 GB
Computer equipment: PC, Celeron CPU 2.6 GHz, 1 GB RAM

Experimental measured results were obtained from the development department of Krystaly, a.s. The measurement output (shown on Fig. 3) describes the phase and shift phase dependance on the excitation frequency, with marked resonant frequencies.

3.1. Results

Table 2 shows the comparison between measured and computed three dominant resonant frequencies of thickness-shear vibration. The average deviation of computed results form the measurement is rather high, about 15%, but the relative distances between particular frequencies are well-kept.

[1]The spatial reduction to the x-direction was used.

Graph of electromechanical coupling coefficients, from which the dominant frequencies can be located easily, is shown on the Fig. 5.

On the next several pages, the visualized experimental results are compared to the ideal theoretical state. The Fig. 6 shows the ideal state: active vibrational sections, amplitudes distribution on axis y at the upper plate of the resonator and vertical cut. Aside from this, the Figures 7–11 show the measured results: whole resonator with displayed amplitudes, schematic amplitudes distribution on axis y at the upper plate of the resonator and vertical cut.

The presented results were obtained from the computation with the parameters mentioned above. For finer meshes (leading to larger dimension of the eigenvalue problem), the computational time considerably increases, which is caused by the limited memory capacity. This fact can be partially avoided e.g. by several repetitions of computations for a smaller number of eigenvalues.

The computed modes contain certain amount of the computational noise (mostly evident in the third dominant mode), but their types can be uniquely identified with the ideal states.

Possible use of the results in the design process.: The model allows to follow up the resonators behavior depending on the changes in its shape properties. As an example, the graph on the Fig. 4 shows the dependance of the resonant frequencies on the change of the resonator thickness and the separation of the particular frequencies. The dependance of the resonant frequencies on the thickness should be linear. This fact is well represented by the model.

Such results can be used in the design process of the resonator, e.g. in the optimization process, when we look for such shape of the resonator, that would have the large distances between particular dominant resonant frequencies.

4. Conclusion

The computer implementation of the model brings the comprehensive programme module, which is suitable for determination of the dominant resonant frequencies in real oscillation problems, including the various shaped resonators.

The previous testing confirmed the applicability of the model, with computed results corresponding to the analytic solutions. The model, applied to the real task of the thickness–shear oscillation of the plan–parallel quartz resonator, brings qualitatively demonstrable agreement with the measurement and keeps the frequency separation in the right proportions. The shift between computed and measured dominant resonant frequencies shows the limitations of the model, but after the calibration to a reference instant, the implementation of the model can be used for practical computations. The model embodies accurate response to the change of the input parameters (e.g., it records the linear dependance of the resonant frequencies on the resonator's thickness).

There are some areas of interest for the future work. The main incoming task is to develop the optimization programme module, which would be suitable for the use in shape design process of new resonators. This involves to repeat the solution of the eigenvalue problems many times. With this objective, the need of a computer implementation running faster than the present–day one may arise – e.g. with the help of parallel computation.

Towards the quickness of computation stays the need of very fine meshes to express the complicated oscillation modes. This fact leads to algebraic problems with very large dimensions. The numerical algebraic methods, used in the model, have a good computational performance, but are limited by the memory size on the PC. The possible way to solve more extensive algebraic problems with keeping the good computational performance is using the implementation of multilevel eigenvalue methods.

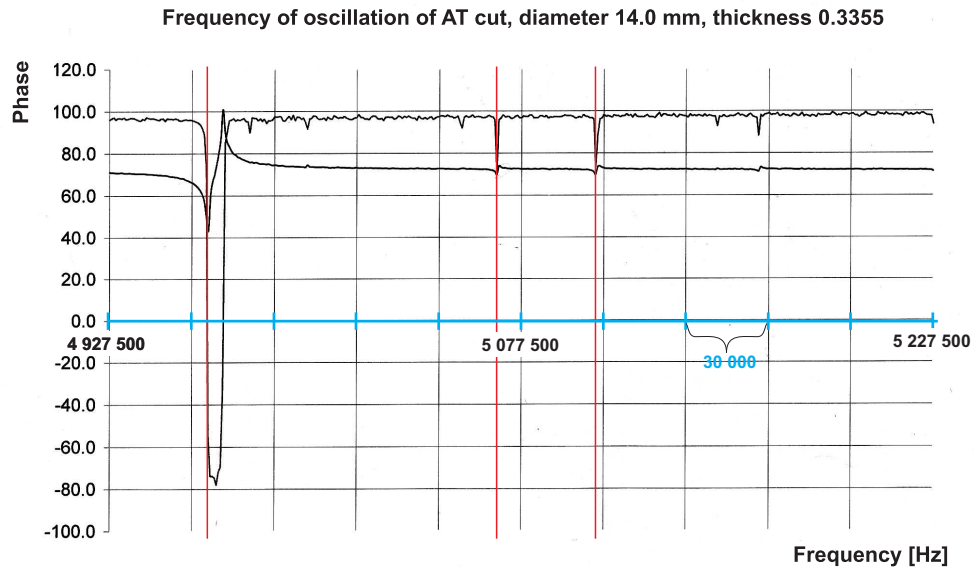


Figure 3: Phase dependance (upper line) and shift phase dependance (bottom line) on the excitation frequency. The maximal phases correspond to the resonances.

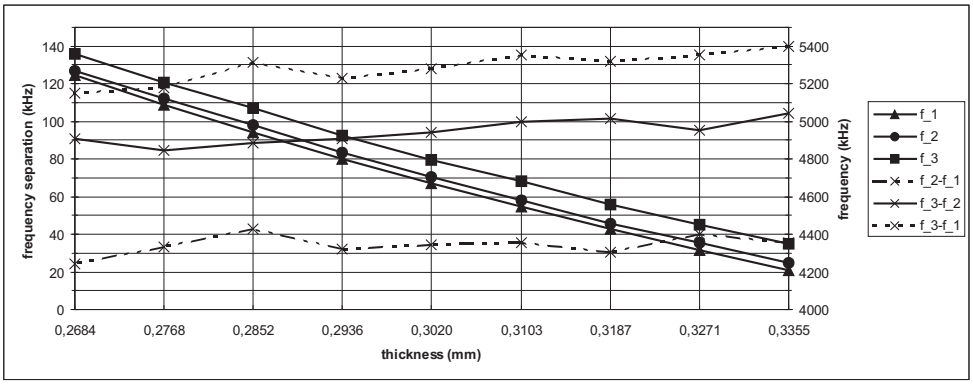


Figure 4: Graph of dependance of resonant frequencies on the thickness of the resonator and intervals between particular frequencies.

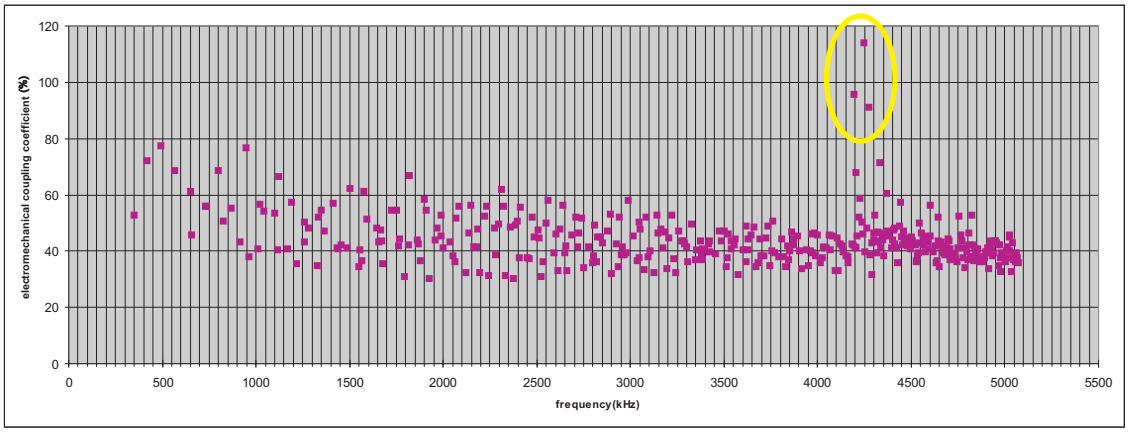


Figure 5: Graph of electromechanical coupling coefficients for first 400 resonant frequencies with marked three dominant oscillation modes.

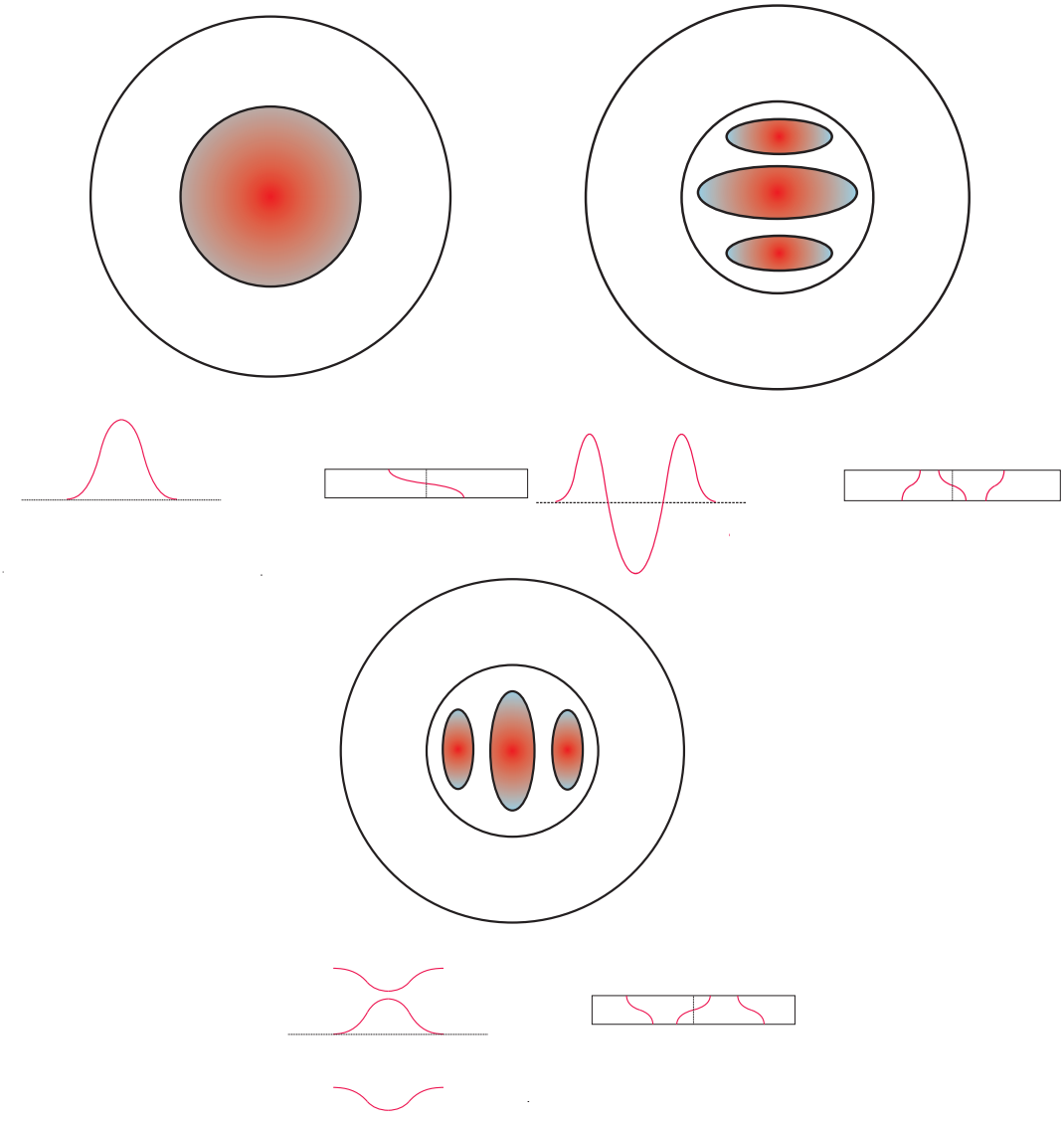


Figure 6: Schematic draft of theoretical shape of the first three dominant modes and distribution of the amplitudes (in the x-direction) on the axis y (left curve) and in the yz-plane median cut.

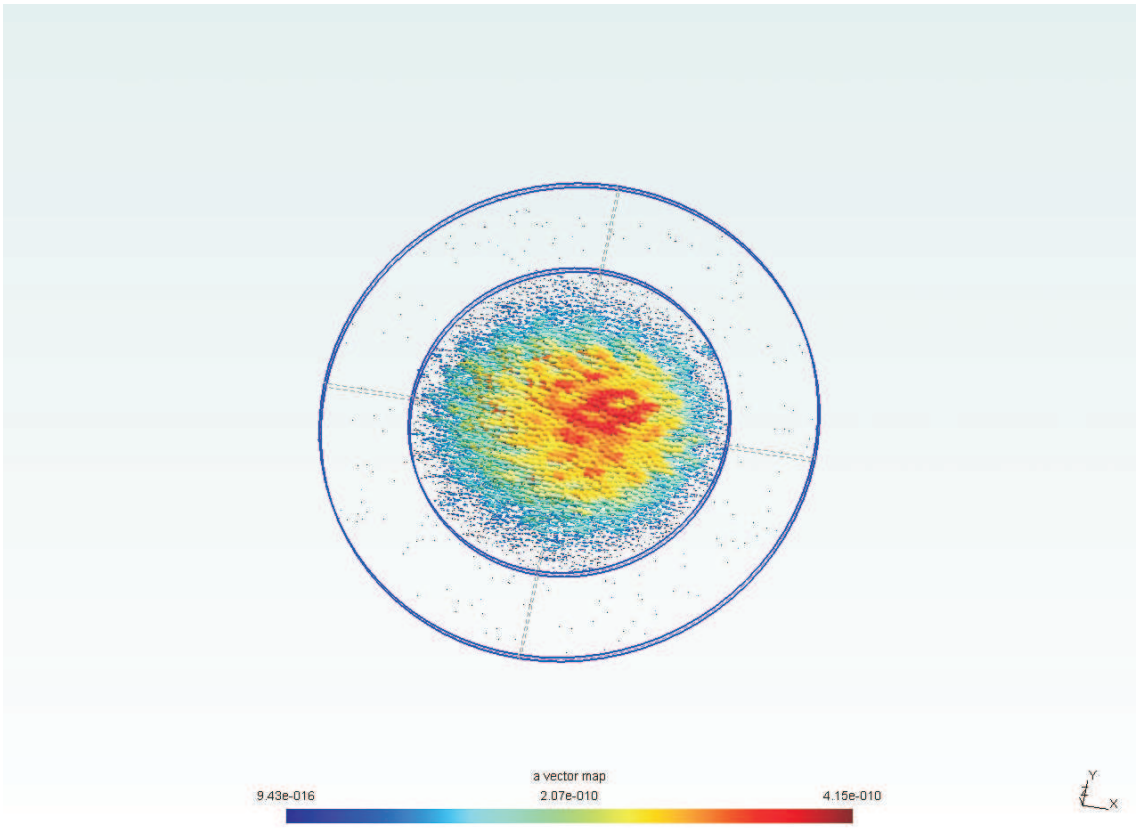


Figure 7: Computed first dominant mode.

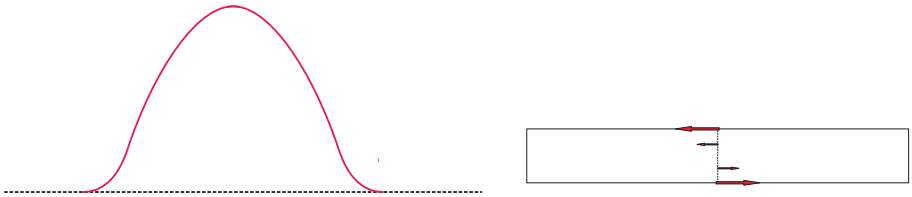


Figure 8: Curve of the amplitudes (in the x-direction) on the axis y (left curve) and in the yz-plane median cut.

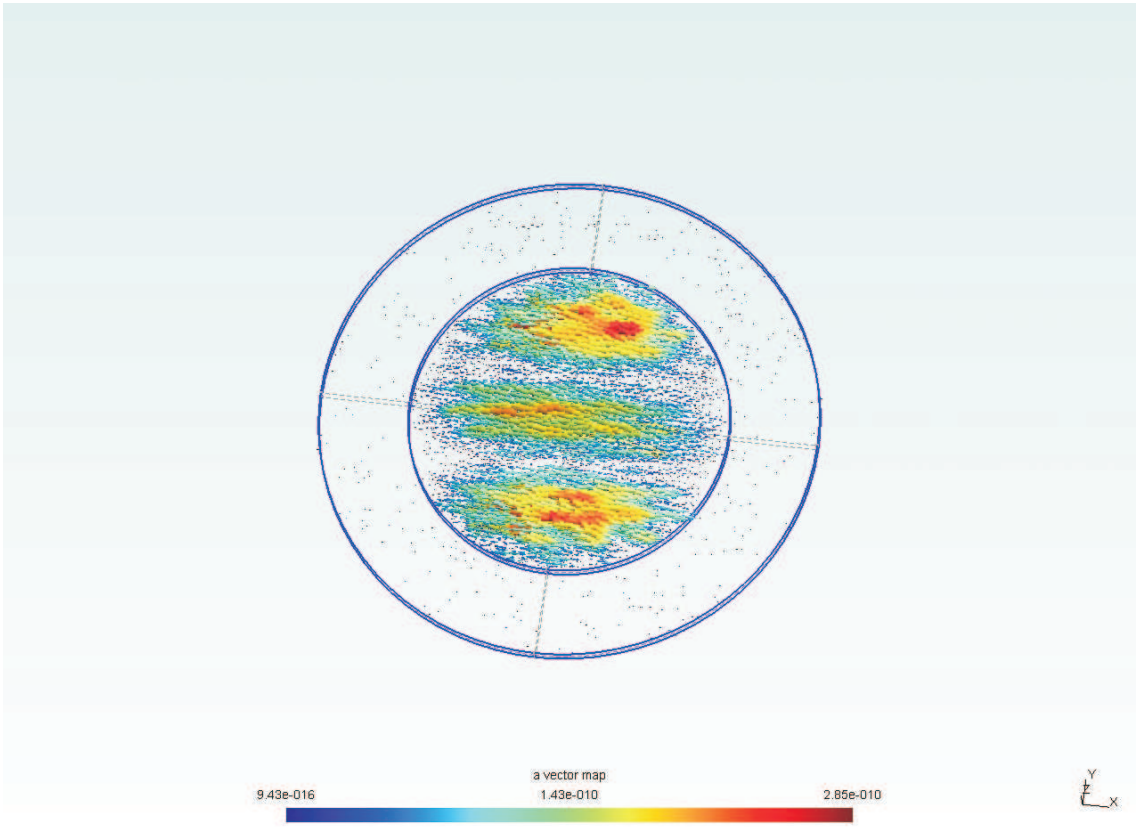


Figure 9: Computed second dominant mode.

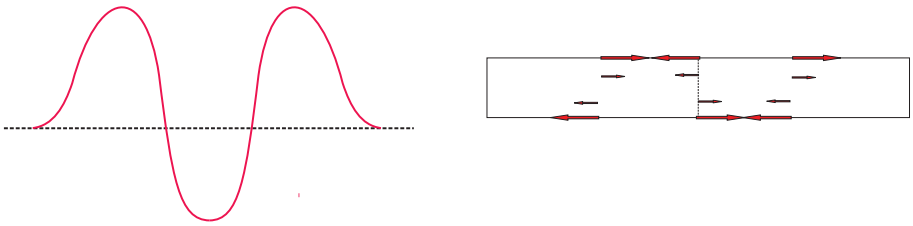


Figure 10: Curve of the amplitudes (in the x-direction) on the axis y (left curve) and in the yz-plane median cut.

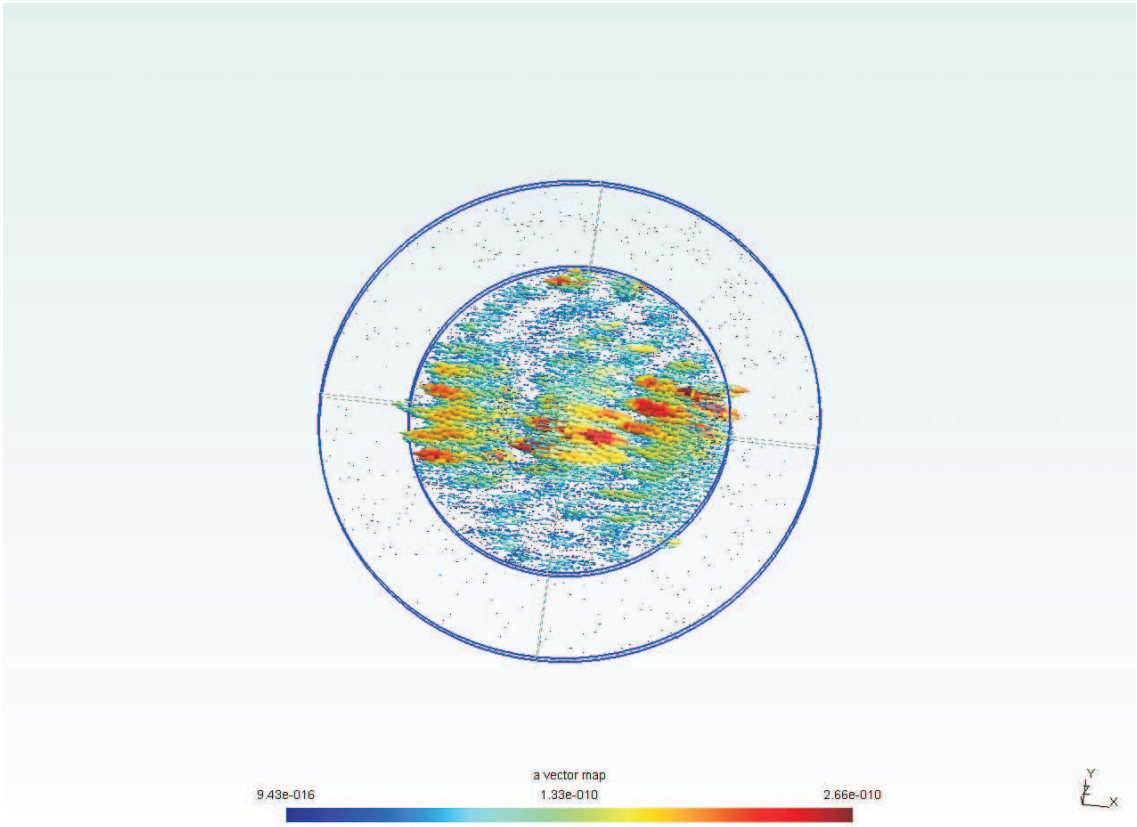


Figure 11: Computed third dominant mode.

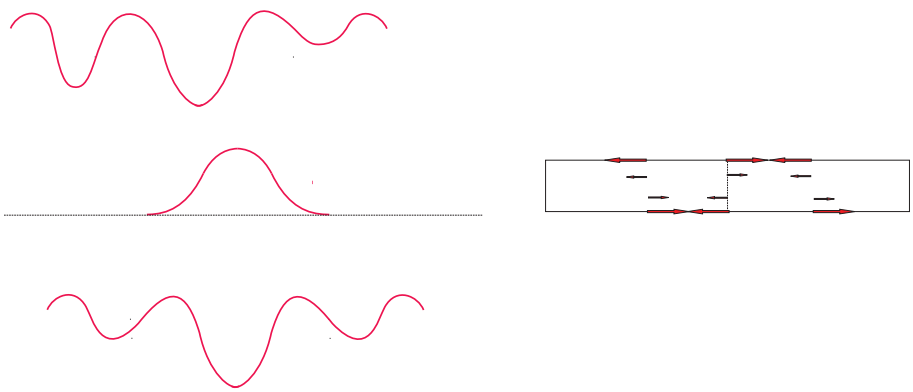


Figure 12: Curve of the amplitudes (in the x-direction) on the axis y (left curve) and in the yz-plane median cut.

References

[1] Lerch R., "Simulation of Piezoelectric Devices by Two- and Three-Dimensional Finite Elements", *IEEE Trans. on Ultrason., Ferroel. and Frequency Control*, Vol. 37, No. 2 (1990), pp. 233- 247.

[2] Rálek P., "Modelling of piezoelectric materials", *Proceedings of the IX. PhD. Conference ICS, Academy of Sciences of the Czech Republic*, Prague 2004.

[3] Rálek P., "Modelling of piezoelectric materials", *Proceedings of the X. PhD. Conference ICS, Academy of Sciences of the Czech Republic*, Prague 2005.

[4] R. Lehoucq, K. Maschhoff, D. Sorensen, Ch. Yang: `www.caam.rice.edu/software/ARPACK/`

[5] Ch. Geuzaine, J.-F. Remacle: `http://www.geuz.org/gmsh/`

[6] O. Marques: `http://crd.lbl.gov/õsni/#Software`

[7] `http://www.krystaly.cz`

Odhadování struktury a asociativní úložiště dat

doktorand:
ING. MARTIN ŘIMNÁČ

Ústav informatiky AV ČR
Pod Vodárenskou věží 2
182 07 Praha 8

Fakulta elektrotechnická
České vysoké učení technické
Technická 2

166 27 Praha 6 - Dejvice

rimnacm@cs.cas.cz

školitel:
ING. JÚLIUS ŠTULLER, CSC.

Ústav informatiky AV ČR
Pod Vodárenskou věží 2

182 07 Praha 8

stuller@cs.cas.cz

obor studia:
Databázové systémy

Práce byla podpořena projektem 1ET100300419 programu Informační společnost "Inteligentní modely, algoritmy, metody a nástroje pro vytváření sémantického webu" a částečně i výzkumným záměrem AV0Z10300504 "Informatika pro informační společnost: Modely, algoritmy, aplikace".

Abstrakt

Odhad struktury dat získaných například z webových zdrojů lze využít jednak pro uložení dat, tak pro netriviální dotazování nad těmito daty. Článek rozšiřuje metodu odhadu struktury dat získávající odpovídající schéma relačního modelu ze vstupních dat a popisuje metodu uložení dat pomocí jednoduchého asociativního úložiště dat právě na základě odhadnutého modelu. Článek diskutuje dvě možné implementace úložiště: první uchovávající data jako instance funkčních závislostí, druhou uchovávající pouze instance funkčních závislosti mezi jednoduchými atributy rozšířenou o podporu komplexních atributů pomocí metainformace.

1. Úvod

Struktura dat představuje soubor všeobecně platných vztahů nad danou množinou dat, což následně vede k popisu každého prvku z této množiny jednotným způsobem na základě zvoleného modelu dat. Tento model může například vycházet z teorie relačních databází, přičemž vztahy mezi atributy jsou popisovány pomocí funkčních závislostí.
Odpovídající množinu funkčních závislostí lze získat postupným testováním, např. pomocí naivního algoritmu, přičemž se testuje vztah mezi každým prvkem z potenční množiny atributů a jedním konkrétním atributem, což vede na algoritmus s nepolynomiální složitostí.
Odpovídající data mohou být následně uložena jako množina instancí nalezených funkčních závislostí, navrhované úložiště tedy musí umožnit pokrýt i instance funkčních závislostí s komplexním atributem, což vede na komplikovanější mechanismus práce s úložištěm.

Článek popisuje návrh zjednodušení výše nastíněného úložiště inspirované formátem RDF (Resource Description Framework) [1], který je používaný v prostředí sémantického webu [2]. Výhodou tohoto formátu je jeho jednoduchost a vysoká popisovací schopnost; data jsou uložena ve formě uspořádaných trojic $< objekt, vlastnost, hodnota >$, nebo alternativně pomocí trojic $< objekt, vztah, objekt >$, odpovídající binárnímu predikátu. Data uložená v tomto formátu ze své podstaty nemusejí být funkčního charakteru.

Podobné úložiště založené na popisu funkčními závislostmi bude organizováno pomocí asociačních pravidel mezi dvěma atomickými entitami včetně pojmenování tohoto vztahu. Takové úložiště bude dále pomocí metainformací rozšířeno o podporu popisu vztahů mezi komplexními atributy bez nutnosti ukládat i odpovídající instance, což vede na redukci paměťových nároků úložiště.

Článek mimo jiné ukazuje, že takové řešení je možné použít za předpokladu, že vstupními data již obsahují nebo budou rozšířena o atomický atribut, který představuje primární klíč všech objektů.

2. Odhadování struktury dat

Motivací pro `odhadování struktury dat` [3] je efektivní uložení dat získaných pomocí specializovaných nástrojů [4, 5] automaticky extrahujících data z webové stránky. Vhodným zdrojem dat pro tyto metody bývají takzvané `produktové katalogy`; webové stránky prezentující technické informace, přičemž formát většinou odpovídá nějakému implicitnímu popisu dat.

Připomeňme, že úloha odhadu struktury dat má význam pouze na `extensionální úrovni`, tedy výsledek úlohy je vždy svázán s konkrétními vstupními daty C (extensionální úroveň výsledku je značena horním indexem, např. M^C). Naopak, definování úlohy na intesionální úrovni je principielně nemožné.

Různé metody pro odhad struktury dat jsou založeny na `naivním algoritmu` uvedeném například v [7]: algoritmus postupně prochází všechny prvky z množiny atributů $A \in \mathscr{A}^C$ (cyklus 1) implikovanou vstupními daty C a pro každý prvek z potenční množiny atributů $X \in \mathscr{P}(\mathscr{A}^C - \{A\})$ (cyklus 2) testuje existenci funkčního zobrazení $f : \mathscr{D}_\alpha^C(X) \rightarrow \mathscr{D}_\alpha^C(A)$ (test 3), přičemž symbol $\mathscr{D}_\alpha^C(A)$ značí aktivní doménu atributu A vzhledem k C. Na základě existence této funkce (na extensionální úrovni vzhledem k danému C) predikuje možnost existence funkční závislosti $(X \rightarrow A)$ na úrovni intensionální.

$$
\begin{aligned}
&M^C = \emptyset \\
&\text{for } \forall A \in \mathscr{A}^C && (1)\\
&\quad \text{for } \forall X \in \mathscr{P}(\mathscr{A}^C - A) && (2)\\
&\quad\quad \text{if } \exists \mathscr{I} : \mathscr{D}_\alpha^C(X) \rightarrow \mathscr{D}_\alpha^C(A) \text{ then} && (3)\\
&\quad\quad\quad M^C := M^C \cup \{X \rightarrow A\}
\end{aligned}
$$

Označíme-li n počet atributů v cyklu (1) a m počet objektů ve vstupní kolekci, složitost tohoto algoritmu je dána převážně velikostí potenční množiny 2^{n-1} (cyklus 2), někdy bývá diskutována [9] i složitost testu (3) funkční závislosti $O((nm) \cdot m)$. Celková složitost naivního algoritmu tedy je:

$$O(n \cdot 2^{n-1} \cdot nm^2) = O(2^{n-1}n^2m^2) \tag{4}$$

Předmětem výzkumu v této oblasti, často označované jako `functional dependency discovery` [3, 6, 7, 8, 9, 10, 11], je redukce počtu testovaných funkčních závislostí. Formálně hledáme množinu $P' \subset P = \mathscr{P}(\mathscr{A}^C - A)$, která bude implikovat shodný model jako původní P. Převážně se jedná o zavedení postupné dekompozice vedoucí k modelu bez triviálních funkčních závislostí M_T.

$$A_{i1} \rightarrow A_j \in M \rightsquigarrow A_{i2} \rightarrow A_j \in M_T \;\; \forall A_{i2} \supset A_{i1} \tag{5}$$

Někdy bývá zohledněno i hledání minimální množiny [3] funkčních závislostí $M_S \subset M$ implikující stejný model díky transitivitě funkčních závislostí.

$$A_i \rightarrow A_j \in M_S \wedge A_j \rightarrow A_k \in M_S \rightsquigarrow A_i \rightarrow A_k \in M \tag{6}$$

3. Asociativní úložiště dat

Předpokládejme, že pomocí zvolené metody odhadneme strukturu dat na základě vstupních dat C. Diskutujme nyní nejjednodušší možný způsob, jak data uložit [12].

Pro začátek neuvažujme funkční závislosti s komplexními atributy. Za tohoto předpokladu množina možných levých stran funkčních závislostí se redukuje pouze na $P' = \mathscr{A}^C$. Díky tomuto omezení

získáváme polynomiálně složitý algoritmus pro odhad modelu, navíc bez nutnosti testování funkčních závislostí s komplexními atributy.

$$O((n \cdot n) \cdot (m)m) = O(n^2 m^2) \tag{7}$$

Navrhněme úložiště Φ popisující instance funkčních závislostí M^{Φ}. Základním modelovacím prvkem zvolíme `element` představující dvojici atribut-hodnota $(A; a)$. Na základě vstupních dat přiřadíme množinu elementů:

$$E = \{e = (A; a) : \quad A \in \mathscr{A}^C, a \in \mathscr{D}_{\alpha}^C(A)\} \tag{8}$$

$$|E| \leq nm \tag{9}$$

Úložiště představuje množinu instancí funkčních závislostí, na něž můžeme nahlížet jako na asociativní pravidla mezi elementy

$$\Phi = \{e_i \rightarrow e_j : e_i, e_j \in E\} \tag{10}$$

Tato representace je nápadně podobná organizaci RDF dokumentů (pro $e_i = (A_i; a_{i'}), e_j = (A_j; a_{j'})$) a je na ni převeditelná.

$$\begin{array}{l} < A_i \qquad \text{rdf:about='}a_{i'}\text{'} > \\ \qquad\quad < A_j \text{ rdf:resource='}a_{j'}\text{'}/ > \\ < /A_i > \end{array} \tag{11}$$

Úložiště Φ je `konzistentní`, pokud

$$\forall a_{j'1} \forall a_{j'1} \nexists a_{j'2} \neq a_{j'1} : (A_i; a_{i'}) \rightarrow (A_j; a_{j'1}) \in \Phi \wedge (A_i; a_{i'}) \rightarrow (A_j; a_{j'2}) \in \Phi \tag{12}$$

Analogicky množina elementů $X \subseteq E$ je konzistentní, pokud

$$\forall a_{j'1} \forall a_{j'1} \nexists a_{j'2} \neq a_{j'1} : (A_j; a_{j'1}) \in X \wedge (A_j; a_{j'2}) \in X \tag{13}$$

Pokud úložiště Φ obsahuje pouze instance funkčních závislostí M^{Φ}, je vždy konzistentní. Pak je možné z úložiště odvodit strukturu dat na základě pravidla:

$$\forall A_i, A_j \in \mathscr{A}^C \; \exists a_{i'} \in \mathscr{D}_{\alpha}^C(A_i), a_{j'} \in \mathscr{D}_{\alpha}^C(A_j) : (A_i; a_{i'}) \rightarrow (A_j; a_{j'}) \in \Phi \rightsquigarrow A_i \rightarrow A_j \in M^{\Phi} \tag{14}$$

Zpětně, struktura dat implikuje pozice v úložišti, které mohou nést informaci.

$$\forall A_i, A_j \in \mathscr{A}^C : A_i \rightarrow A_j \notin M^{\Phi} \rightsquigarrow \forall a_{i'} \in \mathscr{D}_{\alpha}^C(A_i), \forall a_{j'} \in \mathscr{D}_{\alpha}^C(A_j) : (A_i; a_{i'}) \rightarrow (A_j; a_{j'}) \notin \Phi \tag{15}$$

Ze vztahu (14) vyplývá, že není nutné externě udržovat informaci o množině funkčních závislostí platných na úložišti Φ, neboť ji lze získat nejvýše v $O(n^2 m^2)$.

3.1. Odvozování - generalizace

Nadefinujme nyní generalizační odvozovací mechanismus jako zobrazení G:

$$G(\Phi) : E \rightarrow E \tag{16}$$

Toto zobrazení $G(\Phi)$ na zadanou množinu elementů (ve smyslu omezujících podmínek dotazu) $X \subseteq E$ vrátí množinu elementů $Y \subseteq E$, které jsou podmínkami na základě Φ implikovány. Pro úložiště respektující popis pomocí množiny funkčních závislostí platí:

$$X \subseteq Y \tag{17}$$

Vztah (17) platí díky tomu, že úložiště pokrývá i triviální funkční závislosti typu $A \rightarrow A \; \forall A \in \mathscr{A}^C$ (definice viz (5)). Důsledkem je fakt, že mechanismus generalizace je v takovém případě monotónní.

Odvozovací mechanismus je založen na pravidle

$$(A_i; a_{i'}) \in X \wedge (A_i; a_{i'}) \rightarrow (A_j; a_{j'}) \in \Phi \rightsquigarrow (A_j; a_{j'}) \in Y \tag{18}$$

Odvozování dalších (implikovaných) faktů pomocí pravidla (18) z X je realizováno přes vlastnost tranzitivity (6) funkčních závislostí.

Protože úložiště pokrývá pouze instance funkčních závislostí (14), je zaručena konzistence úložiště (12). Generalizační mechanismus zaručuje, že na konzistentní vstup (13) bude vrácen konzistentní výstup.

Toto tvrzení vychází z předpokladu, že množina X je konzistentní. Neboť dle (17) $X \subseteq Y$, pak část odpovídající X je rovněž konzistentní. Stačí tedy ukázat konzistenci množiny $Y - X$. Aby došlo k aktivaci nějakého dalšího y na základě pravidla (18) aplikovaném na nějaké konkrétní $x \in X$, díky konzistenci úložiště Φ odvozené y musí být rovněž konzistentní vůči všem elementům v Y.

4. Komplexní atributy a úložiště

Uvažujeme funkční závislost atributu A_r na komplexním atributu $\{A_{l_1} \dots A_{l_{n'}}\}$:

$$\{A_{l_1} \dots A_{l_{n'}}\} \rightarrow A_r \in M \tag{19}$$

Diskutujme nyní dvě varianty vyjádření vztahu mezi komplexními atributy (funkční závislost s komplexním atributem na levé straně).

- Uložení instance funkční závislosti s komplexním atributem.
- Uložení metainformace o existenci funkční závislosti s komplexním atributem.

4.1. Komplexní atributy v úložišti

Tato varianta předpokládá, že úložiště je schopno pojmout informaci o komplexním atributu a jeho instanci. Prakticky takové úložiště lze implementovat jako množinu uspořádaných trojic

$$\{< id, \{(A_{l_1}; a_{l_1}) \dots (A_{l_{n'}}; a_{l_{n'}})\}, (A_r; a_r) >\} \tag{20}$$

kde id je identifikátor instance, $\{(A_{l_i}; a_{l_i})\} \subseteq \mathscr{P}(E)$ je hodnota komplexního atributu (tj. hodnoty representující podmínku aktivace instance komplexního atributu) a $(A_r; a_r) \in E$ je fakt reprezentující důsledek aktivace.

Na takové úložiště implementované pomocí atomických atributů (tj. splňující požadavky 1NF) bude potřeba $n' + 1$ instancí dvojic mezi identifikátorem id a elementy $(A_i; a_{i'}) \forall i \in \{l_1 \dots l_{n'}\} \cup \{r\}$.

Funkční závislost mezi jednoduchými atributy je (jako triviální důsledek) namapována na 2 dvojice.

4.2. Komplexní atribut jako metainformace

Způsob uvedený v předchozí sekci vede na poměrně komplikovaně organizované úložiště. Pokusme se tedy navrhnout jednodušší způsob. Modelování funkční závislosti s komplexním atributem (19) se rozpadá na dva případy. Pokud:

- existuje atribut A_k takový, že $A_k \rightarrow A_{l_i} \in M^\Phi$ pro $\forall i = 1 \dots n'$, pak taková závislost bude v asociativním úložišti Φ interpretována jako soubor instancí funkčních závislostí s atributem A_k na levé straně a atributy A_{l_i} a A_r na straně pravé. Tedy:

  $$\forall i = 1 \dots n' \; \exists A_k : A_k \rightarrow A_{l_i} \in M \rightsquigarrow A_k \rightarrow A_{l_i} \in M^\Phi \wedge A_k \rightarrow A_r \in M^\Phi \tag{21}$$

  Informace o existenci funkční závislosti s komplexním atributem v M je zachycena jako metainformace pomocí faktu

  $$\{A_{l_1} \dots A_{l_n}\} \rightarrow A_r \in M^\Phi_\odot \tag{22}$$

- takový atribut A_k neexistuje, asociativní úložiště nemá žádný prostředek, jak svázat hodnotu komplexního atributu $\{A_{l_i}\}$ na levé straně s hodnotou atributu A_r na pravé straně a tedy tento vztah nebude úložištěm pokryt.

Existence atributu A_k je tedy podmínkou nutnou, aby vůbec došlo k zachycení vztahu mezi atributy. V praxi tuto podmínku triviálně zaručíme, pokud každému záznamu ve vstupní kolekci přiřadíme jednoznačný identifikátor A_I. Z definice funkční závislosti vyplývá, že $A_I \rightarrow A$ pro $\forall A \in \mathscr{A}^C$, tedy druhá varianta nemůže nastat, neboť v takovém případě minimálně $A_k = A_I$.

4.3. Instance funkčních závislostí s komplexním atributem

Diskutujme nyní úpravy asociativního úložiště (20) ve smyslu (21) tak, aby bylo možné rozšířit generalizační mechanizmus (17 -18) o podporu komplexních atributů na základě metainformace v $M^{\Phi}_{\odot}$.

Předpokládejme, že podmínka existence klíče (21) je splněna. Ukažme, že za tohoto předpokladu není potřeba ukládat instance funkční závislosti s komplexním atributem (19), neboť generalizační krok lze provést na základě znalosti hodnoty/hodnot společného atributu A_k, která je odvozena "zpětně" z hodnoty komplexního atributu.

$$\{A_{l_1} \dots A_{l'_n}\} \rightarrow A_r \in M^{\Phi}_{\odot} \wedge \forall (A_{l_i}; a_{l_i}) \in X \rightsquigarrow (A_r; a_r) \in Y \tag{23}$$

Zaveďme aktivaci komplexního atributu. Úložiště obsahuje instance

$$(A_k; a_k) \rightarrow (A_{l_i}; a_{l_i}) \in \Phi \qquad i = 1 \dots n' \tag{24}$$

$$(A_k; a_k) \rightarrow (A_r; a_r) \in \Phi \tag{25}$$

Generalizační krok spočívá v nalezení množiny možných elementů $(A_k; a_k)$ relevatních k $(A_r; a_r)$ tak, že $a_k \in K(X, A_k)$, přičemž

$$K(X, A_k) = \{a_k : \left\{ \begin{array}{l} \{A_{l_1} \dots A_{l'_n}\} \rightarrow A_r \in M^{\Phi}_{\odot} \wedge A_k \rightarrow A_r \in M^{\Phi} \\ \forall i = 1 \dots n' : (A_{l_i}; a_{l_{i'}}) \in X \wedge (A_k; a_k) \rightarrow (A_{l_i}; a_{l_{i'}}) \in \Phi \end{array} \right. \} \tag{26}$$

Nyní se generalizační mechanismus rozpadá na tři případy. Pokud:

- $|K(X, A_k)| = 0$, pak není funkční závislost $\{A_{l_1} \dots A_{l'_n}\} \rightarrow A_r$ pokryta instancemi buďto v X nebo v úložišti Φ a tudíž nemůže být aktivována.
- $|K(X, A_k)| = 1$, neboli $\{A_{l_1} \dots A_{l'_n}\} \rightarrow A_k \in M$, pak hodnotu a_r ze vztahu (23) určíme pomocí instance (25). V tomto případě atributy A_r a A_k mohou splývat.
- $|K(X, A_k)| > 1$, neboli $\{A_{l_1} \dots A_{l'_n}\} \rightarrow A_k \notin M$, tedy předchozí mechanismus nelze použít. Za této situace ovšem $\{A_{l_1} \dots A_{l'_n}\} \rightarrow A_r \in M^{\Phi}_{\odot}$ a $A_k \rightarrow A_r \in M^{\Phi}$, pak ale musí nutně $(A_k; a_k) \rightarrow (A_r; a_r) \in \Phi \ \forall a_k \in K(X, A_k)$ vést na jedinou hodnotu a_r (neboť v opačném případě by $\{A_{l_1} \dots A_{l'_n}\} \rightarrow A_r \notin M$).

Tedy, instance funkčních závislosti s komplexním atributem není potřeba ukládat za předpokladu existence nějakého společného klíčového atributu A_k, neboť se podle výše zmíněného postupu dají odvodit z instancí funkčních závislostí mezi jednoduchými atributy.

Pokud bychom místo generalizace dotazu volili jeho specializaci (tj. chtěli bychom získat seznam objektů splňující podmínky dané dotazem), je uvažování komplexních atributů v odvozovacím mechanizmu bezpředmětné, neboť zaručeně dojde k aktivaci atributu A_k, který je vždy specializovanějším nežli libovolný z atributů A_{l_i}. V okamžiku, kdy budeme chtít pomocí generalizace atributů určujících objekty získat jejich popis (tj. konkrétní množiny dvojic $(A_i; a_{i'})$), dojde nejprve k aktivaci atributu A_k a v následném kroku pak k aktivacím atributů A_{l_i} i A_r, tj. v tomto případě není nutné uvažovat rozšíření (23) generalizačního mechanismu.

Naopak rozšíření (23) je nutné uvažovat v případě, že X obsahuje instance všech atributů A_{l_i}, avšak neobsahuje aktivaci A_k, což v praxi nastává poměrně často.

4.4. Úložiště s metainformacemi

Diskutujme nyní možnost, jak implementovat úložiště s vlastnostmi odpovídajícími předchozím dvěma odstavcům.

Neboť M^{Φ} obsahuje funkční závislosti pouze mezi jednoduchými atributy, není nutné uchovávat identifikátor instance, postačuje přímo uchovávat jen dvojice elementů $(A_i; a_{i'}) \rightarrow (A_j; a_{j'}) \in \Phi$ například pomocí množiny uspořádaných dvojic

$$\{< (A_i; a_{i'}), (A_j; a_{j'}) >\} \tag{27}$$

Funkční závislosti s komplexními atributy jsou uloženy v $M^{\Phi}_{\odot}$, způsob uložení je analogický k (20) s tím rozdílem, že se uchovává informace pouze o funkční závislosti, nikoliv o jejích instancích.

$$\{< id, \{A_{l_1} \dots A_{l_{n'}}\}, A_j >\} \tag{28}$$

4.5. Případová studie - porovnání

Porovnejme závěrem obě navržené implementace úložiště podle počtu záznamů nutných k uložení informace, první popisující instance potencionálně komplexních atributů a druhý, obsahující pouze instance funkčních závislostí s jednoduchými atributy. Pro jednoduchost uvažujme, že již na začátku ukládaní máme k dispozici model popisující vstupní kolekci dat.

Pro případovou studii byla zvolena data reprezentující n'-ární součet hodnot atributů $A_1 \dots A_{n'}$, hodnota součtu je pak uložena pod atributem A_S. Každý záznam je rozšířen o primární klíč A_k z důvodů uvedených u (21).

$$\begin{array}{ccccccc}
A_k & A_1 & A_2 & \cdots & A_{n'-1} & A_{n'} & A_S \\
0 & 0 & 0 & & 0 & 0 & 0 \\
1 & 0 & 0 & & 0 & 1 & 1 \\
2 & 0 & 0 & & 0 & 2 & 2 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
9 & 0 & 0 & & 0 & 9 & 9 \\
10 & 0 & 0 & & 1 & 0 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
19 & 0 & 0 & & 1 & 9 & 10 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
10^{n'} & 9 & 9 & & 9 & 9 & 9n'
\end{array} \tag{29}$$

Model implikovaný vstupní kolekcí dat je

$$A_k \rightarrow A_S \tag{30}$$

$$A_k \rightarrow A_i \ \forall i = 1 \dots n' \tag{31}$$

$$\{A_1, \dots, A_n\} \rightarrow A_S \tag{32}$$

Každý záznam ze vstupních dat (29) implikuje

- 1 instanci funkční závislosti (30),
- n' instancí funkčních závislostí (31) a
- 1 instanci funkční závislosti (32) s komplexním atributem.

Tedy počet instancí se během jednoho kroku změní o:

$$\Delta|I^B| = n' + 1 \tag{33}$$
$$\Delta|I^C| = 1 \tag{34}$$

Pokud budeme uvažovat uložení předmětného záznamu ze vstupní dat do úložiště pokrývající přímo instance komplexních atributů pomocí (20), je potřeba na uložení jedné instance funkční závislosti s komplexním atributem arity a na levé straně potřeba celkem $(a+1)$ záznamů. V tomto konkrétním případě (29) se interní struktury úložiště rozšíří o m záznamů, přičemž

$$m = \Delta|I^B| \cdot 2 + \Delta|I^C| \cdot (n' - 1 + 1) = 3n' + 2 \tag{35}$$

Budeme-li uvažovat úložiště pokrývající pouze instance funkčních závislostí s jednoduchými atributy získaných podle (21) uložených jako (28), každou takovou instanci pokryje 1 záznam v interní struktuře úložiště, tedy přírůstek m' pro tento případ je

$$m' = \Delta|I^B| \cdot 1 = n' + 1 \tag{36}$$

Neboť se jako metainformace ukládá seznam funkčních závislostí s komplexními atributy, zůstává velikost $|M_{\odot}^{\Phi}|$ nezměněna.

Srovnáním (35) a (36) získáme

$$\frac{m'}{m} = \frac{n' + 1}{3n' + 2} \simeq 33\% \tag{37}$$

5. Závěr

Článek se zabývá uložením vstupních dat do úložiště pomocí odhadnuté struktury. Ukazuje, že intuitivně nejjednodušší možný návrh úložiště Φ jako asociativní paměti $(A_i; a_{i'}) \rightarrow (A_j; a_{j'})$ primárně popisující vztahy mezi jednoduchými atributy je postačující, neboť obsahuje veškeré informace nutné pro podporu vztahů platných mezi komplexními atributy za předpokladu (21), že existuje nějaký atribut jednoznačně definující každý záznam ve vstupní kolekci dat. Tento fakt vede k nutnosti uložení pouze instancí funkčních závislostí s jednoduchými atributy, přičemž vztahy mezi komplexními atributy jsou vyjádřeny externě pomocí metainformace $M^{\odot}$.

Ukazuje se, že analýzu těchto vztahů je nutné použít pouze při generalizaci, generalizační mechanismus je nutné rozšířit o pravidlo (23).

V článku prezentovaný výsledek vede ke značným paměťovým úsporám, jednoduššímu návrhu úložiště a v neposlední řadě k odstranění případných duplicit a nekonzistenci dat v úložišti. Vzhledem k metodám odhadu dat se jedná o rozdělení výpočetní složitosti algoritmu (4) na dvě části, první polynomiálně složitou pro odhad struktury (7) a uložení dat (bez uvažování komplexních atributů) a druhou, nepolynomiálně složitou generující příslušná metadata v $M_{\odot}^{\Phi}$. Proto tuto část je ale možné přejmout výsledky z [3, 6, 7, 8, 9, 10, 11]

V závěru článku je uvedeno porovnání dvou popisovaných přístupů k úložišti dat, první je implementované s přímou podporou instancí funkčních závislostí s komplexními atributy a druhá bez této podpory. Obě implementace jsou navrženy tak, aby splňovali podmínku 1NF. Díky tomuto faktu je nutné ukládat jednu instanci funkční závislosti s komplexním atributem pomocí několika záznamů. Případová studie v jednom konkrétním případě (n-ární součet čísel) ukazuje, že úložiště bez pokrytí vztahů mezi komplexními atributy vykazuje asi 66% úsporu paměťových nároků úložiště. Jinými slovy případová studie potvrzuje výhodu pokrytí informace pomocí jednoduchých asociačních pravidel a v případě nutnosti uvažování metainformace o vztazích mezi komplexními atributy.

Závěrem poznamenejme, že z korespondujících důvodů nebyla do formátu RDF zavedena podpora komplexních atributů.

Praktické závěry tohoto článku jsou nasnadě, v případě potřeby uložení velkého počtu dat, které lze předpokládat i u extrakce dat z webových stránek, je hledisko paměťových úspor velmi důležité. V tomto konkrétním případě ani neznamená nutnost ztráty informace, ba právě naopak zabraňuje případné nekonzistenci dat vznikající duplicitou mezi instancemi funkčních závislostí s komplexními atributy a odvozovacím mechanismem pro tyto instance.

Literatura

[1] Eric Miller, Ralph Swick, Dan Brickley. "Resource Description Framework". `<http://www.w3.org/RDF/>` [on-line], 2004.

[2] Grigoris Antoniou, Frank van Harmelen. "A Semantic Web Primer". MIT Press, 2004. ISBN: 0-262-01210-3.

[3] M. Řimnáč. Odhad struktury dat a induktivní logické programování In *ITAT 2005: Workshop on Information Technologies, Applications and Theory.* pp. 263-272. ISBN: 80-7097-609-8. 2005

[4] D.W. Ebley, C. Tao, S.W. Liddle. "Automating the extraction of data from HTML tables with unknown structure." In *Data & Knowledge Engineering* 54, pp 3-24. Elsevier. 2005

[5] G. Gottlob, Ch. Koch. " A Formal Comparison of Visual Web Wrapper Generators." In *Proc. of. SOFSEM 2006: Theory and Practice of Computer Science*. LNCS 3831. ISSN 0302-9743. Springer.

[6] P.A. Flach, I. Savnik. "Database Dependency Discovery: A Machine Learnig Approach". In *AI Communications*, Volume 12/3, pp. 139–160. 1999.

[7] H. Mannila, K.J. Räihä. "Dependency Inference". In *Proc. of VLDB*. pp. 155–158. ISBN: 0-934613-46-X. 1987.

[8] H. Mannila, K.J. Räithä. "Algorithms for inferring functional dependencies from relations." In *Data & Knowledge Engineering* 12, pp 83-99. Elsevier. 1994.

[9] J. Kivinen, H. Mannila. "Approximate Inference of Functional Depentencies from Relations". In *Proc. of 4. int. conf. on Database Theory*, Berlin, Germany. pp. 129–149. ISSN: 0304-3975. 1995.

[10] Ch. Giannella, E. Robertson. "On Approximation Measures for Functional Dependencies". In *ADBIS 2002: Advances in databases and information systems*, Elsevier. pp. 483–507. ISSN 0306-4379. 2004.

[11] T. Akutsu, S. Miyano, S. Kuhara. "A Simple Greedy Algorithm for Finding Functional Relations: efficient implementation and average case analysis." Theoretical Computer Science, Volume 292, Issue 2. pp 481–495. ISSN: 0304-3975. 2003.

[12] D. Bednarek, D. Obdrzalek, J. Yaghob, F. Zavoral. "Access Rights Definition and Management in an Information System based on a DataPile Structure". In *ITAT 2004, Workshop on Information Technologies, Application and Theory*, 2004.

Machine Learning Methods for Knowledge Discovery in Medical Data on Atherosclerosis

Post-Graduate Student:
BS JOSÉ IGNACIO SERRANO

Instituto de Automática Industrial
Spanish Council for Scientific Research (CSIC)
Ctra. Campo Real km 0.200 - La Poveda
28500 Arganda del Rey, Madrid, Spain

nachosm@iai.csic.es

Supervisor:
PROF. RNDR. JANA ZVÁROVÁ, DRSC.

Department of Medical Informatics
Instutite of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic

zvarova@euromise.cz

Field of Study:
Knowledge Discovery in Medical Data

Research was partially supported by the Research Plan of ICS AS CR AV0Z10300504 and the "María Bueno" Research Visits Plan of the Spanish Council for Scientific Research together with the support of the Industrial Automation Institute.

Abstract

Machine learning techniques are methods that given a training set of examples infer a model for the categories of the data, so that new (unknown) examples could be assigned to one or more categories by pattern matching within the model. The data from follow-up studies with repeated collection of the same type of data are very suitable for this analysis. Machine learning algorithms belonging to a variety of paradigms have been applied to knowledge discovery on medical data. All the used algorithms belong to the supervised learning paradigm. Several algorithms have been tested, trying to cover most of the kinds of supervised learning. Two kinds of experiments have been carried out. The first is intended to discover associations between attributes. The second kind is intended to test prediction of future disorders. For the experiments in this paper the data used was from the twenty years lasting primary preventive longitudinal study of the risk factors (RF) of atherosclerosis in middle aged men. Study is named STULONG (LONGitudinal STUdy). The results show that some methods predict some disorders better than others, so it is interesting to use all the algorithms at a time and consider the result confidence based upon the known tendency of each method. The machine learning algorithms have been also used in the prediction of death cause, obtaining poor results in this case, maybe due to the small amount of information (entries) of this type in the dataset.

Keywords: knowledge discovery, supervised machine learning, biomedical data mining, risk factors of atherosclerosis

1. Introduction

Machine learning techniques [1] are methods that given a training set of examples infer a model for the categories of the data, so that new (unknown) examples could be assigned to one or more categories by pattern matching within the model.

Machine learning techniques have been applied successfully to a high variety of problems and data for prediction tasks. The main objective is to research how to apply machine learning algorithms to this data in order to discover relationships between attributes and to make predictions that could be useful for decision support. Medical data is a special kind of data, because many different kinds of features are involved in the collections. Moreover, the medical data have several known problems: missing, incorrect and sparse information and temporal data. Machine learning methods are very suitable for this kind of data [2]. There exists several KDD works attempting to deal with large-scale medical information. In [3], authors try to detect type of hepatitis by extracting short sequential patterns from the temporal features. In [4], simple rules are discovered using 4ft-miner (i.e. statistical tables of two rows and two columns), in order to

temporally characterize, by differences, the hepatitis types B and C. Authors in [5] attempt to discover rules of singles boolean features that can be able to predict the liver fibrosis stage. The same application appears in [6] but, in this case, extracted patterns are clustered and then these clusters are assigned to fibrosis stages depending on the covered examples. They also applied this technique to atherosclerosis risk [7]. The data from follow-up studies with the repeated collection of the same type of data are very suitable for this analysis. Additional examples of data mining on biomedical data are presented in [8] and [9]. For the experiments in this paper the data used was from the twenty years lasting primary preventive longitudinal study of the risk factors (RF) of atherosclerosis in middle aged men. The study is named STULONG (LONGitudinal STUdy) [10]. The main target of this study is to validate machine learning as a way of association mining and to validate classification performance as a measurement of the salience for the discovered association. It is also intended to test machine learning algorithms in the prediction of far future disorders.

In the next section, the STULONG dataset is briefly presented. In section 3, the machine learning algorithms tested are described. Section 4 presents the measures for evaluation and Section 5 describes validation experiments. Finally, concluding remarks and future work is presented in Section 6.

2. Description of the Study and Data Set

The STULONG (http://euromise.vse.cz/challenge2004/index.html) [10] [11] [12] data were collected by the 2nd Department of Internal Medicine, 1st Faculty of Medicine and General University Hospital, Prague and transferred to the electronic form and analysed by statistical methods by the European Centre of Medical Informatics, Statistics and Epidemiology of Charles University and the Academy of Science of the Czech Republic, Prague.

The main aims of the study were:

1. To identify the prevalence of risk factors (RF) of atherosclerosis in a population generally considered as the most endangered by possible complications of atherosclerosis, i.e. middle aged men.
2. To follow the development of these RF and their impact on the examined men health, especially with respect to atherosclerotic cardiovascular diseases.
3. To study the impact of complex intervention of RF on their development and cardiovascular morbidity and mortality in men.

Men were divided according to presence of risk factors (RF), overall health conditions and ECG result into three groups: Normal (NG), Risk (RG) and Pathologic (PG). These group will be referred in this paper as level groups [12].

Four data files have been used for the analysis:

1. The file ENTRY contains values of 244 attributes obtained from entry examinations for each man; these attributes are either codes or results of size measurements of different variables or results of their transformations (identification of man, family and personal history, social factors - education, physical activities, smoking, eating habits, alcohol, after them anthropometric measurements - height, weight, skin folds, physical examination with measurement of blood pressure, pulse, laboratory values and coding of ECG).
2. The file CONTROL contains results of observation of 66 attributes recorded during control examinations. There are attributes corresponding to identification, to habit changes, to personal history, physical examination and biochemical values, and data about hypertension, hypercholesterolemia, hypertrigliceridemia and other coronary and oncological diseases. This file consists of 10,572 records of long observation.

3. Additional information about health status of 403 men dropped out in the time of the study was collected by the postal questionnaire. Resulting values of 62 attributes are stored in the LETTER file.
4. There are 5 attributes concerning death of 389 patients. Values of these attributes are stored in the DEATH file. It contains attributes for the identification of the patients and the date and cause of death.

3. Description of the Used Methods

All the used algorithms belong to the supervised learning paradigm. That is, a learning stage is needed in order to build a model over the training examples and then use this model to predict the category of unknown examples. Several algorithms have been tested, trying to cover most of the kinds of supervised learning. Each of the used methods is very briefly explained next:

3.1. Naive Bayes

Naive Bayes [13] calculates, for each pair attribute-value, for example (education, university), the probability of belonging to each category, by dividing the number of examples of the target category where the pair appears by the total number of examples where the pair appears. Thus, each pair will have a probability for each tentative category. Naive Bayes is based on the assumption that every pair attribute-value within an example is independent on each other. Thus, when an unlabeled example is classified, the probability for each category of the example is the multiplication of the probability for the corresponding category of each of the pairs that form the example. The predicted category is the one with the highest probability.

3.2. Multilayer Perceptron

The classification model of the Multilayer Perceptron Neural Network [14] is composed of a certain number of layers of neurons interconnected between them. The architecture used for this dataset is presented in Figure 1.

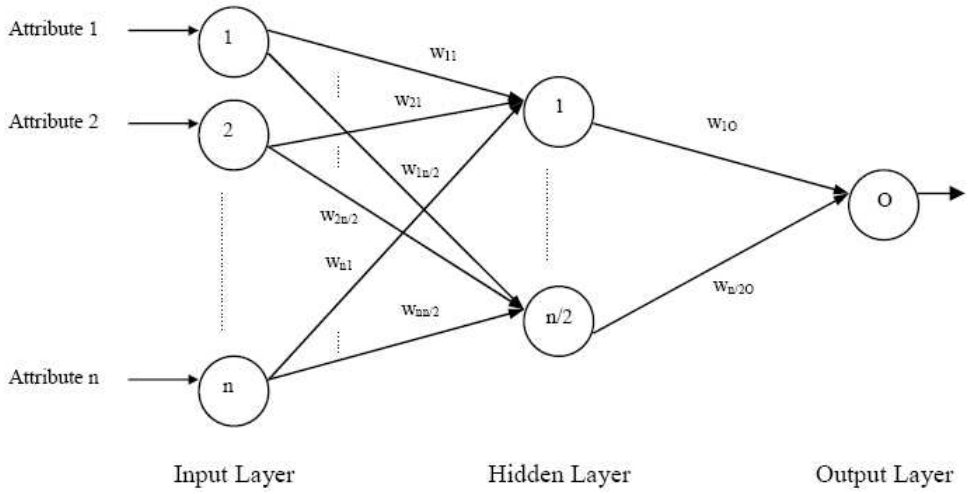


Figure 1: Architecture of the Multilayer Perceptron Neural Network used.

Each connection has an associated weight. The input to each neuron is the weighted sum, using the association weights, of all the incoming values. The output of each neuron is the result of applying a function. In this case, a typical sigmoid function is implemented in all the neurons. Figure 2 shows the function expression and representation.

Thus, each of the attribute values from a sample of the dataset is entered in the corresponding neuron of the input layer, and the values spreads through the network to the output layer, where the output value of the

Figure 2: Expression and representation of sigmoid function.

neuron is the predicted class.

The training phase consists of, given a set of initial weights values, entering each of the labelled examples of the training dataset into the model and comparing the output value with the expected class. Depending on the error of the predicted class, the back propagation algorithm changes weights from the output layer to the input layer, in order to make the predicted value to be more similar to the expected one. This process is carried out a certain number of epochs or iterations. In this case, this number is equal to 500. The amount the weights are changed in back propagation, so called learning rate, is 0.3, and the momentum applied to the weights during updating is 0.2. If the back propagation algorithm does not reach a good approximation to the expected output after one iteration, then it resets the model and causes the learning rate to decrease.

3.3. Support Vector Machines (SVM)

Support Vector Machines [15] try to separate examples, based on their category, in the n-dimensional space, being n the number of attributes or features, by hyper planes of the form w + b, so that

$$xw + b \geq +1 \quad \rightarrow \quad category = true \tag{1}$$
$$xw + b \geq -1 \quad \rightarrow \quad category = false \tag{2}$$

x being the example represented as a vector of n components. Here, w is the support vector, perpendicular to the hyper plane, and correspond to examples that are beyond or over the limits of their category (see Figure 3).

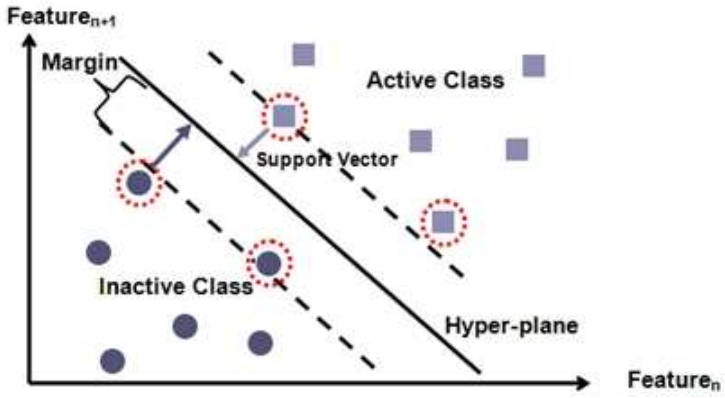


Figure 3: Support vectors scheme.

The support vector also defines, by its module, a margin of one between the hyper plane and the first positive and negative examples (that the reason for +1, -1 thresholds). For each category, the algorithm tries to find w maximizing the margin. To classify an unlabeled example the algorithm simply applies the expression above. This is a simple implementation of the method and the one used in the experiments, but there are other more sophisticated implementations and techniques.

3.4. K-Nearest Neighbour

KNN is a memory-based algorithm [16], with the background idea that past experiences can help us to solve present ones by analogy. It considers each example as a vector of n components, being n the number of attributes or features. It does not need a learning stage. To predict the class of an unlabeled example, the algorithm compares the input example with each of the examples in the training data or memory, by calculating the distance between them. Then, the majority class of the K most similar training examples is the one predicted for the input example. The distance used in the experiments is the Euclidean distance between vectors. However, there are more possibilities in the literature.

3.5. ID3 and C4.5 Decision Trees

The model produced by this algorithm is a tree [17], where each node corresponds to an attribute and each arc of the node corresponds to a possible value of the node attribute.

The learning algorithm constructs the tree from the training data. The selection of the attribute that will form a node, at each moment, is carried out by calculating the entropy of the data after the selection of the node. That is, for each attribute, the entropy of the remaining data without the attribute, separated by the different values of the node attribute, is calculated. Thus, the attribute that produces the minimum entropy is the selected for the node. The process goes on until there is no more attributes or the number of remaining examples under a node is lower than a certain threshold. In the former case, the majority class of these remaining examples is the one settled under the node. In Figure 4, we can see an example:

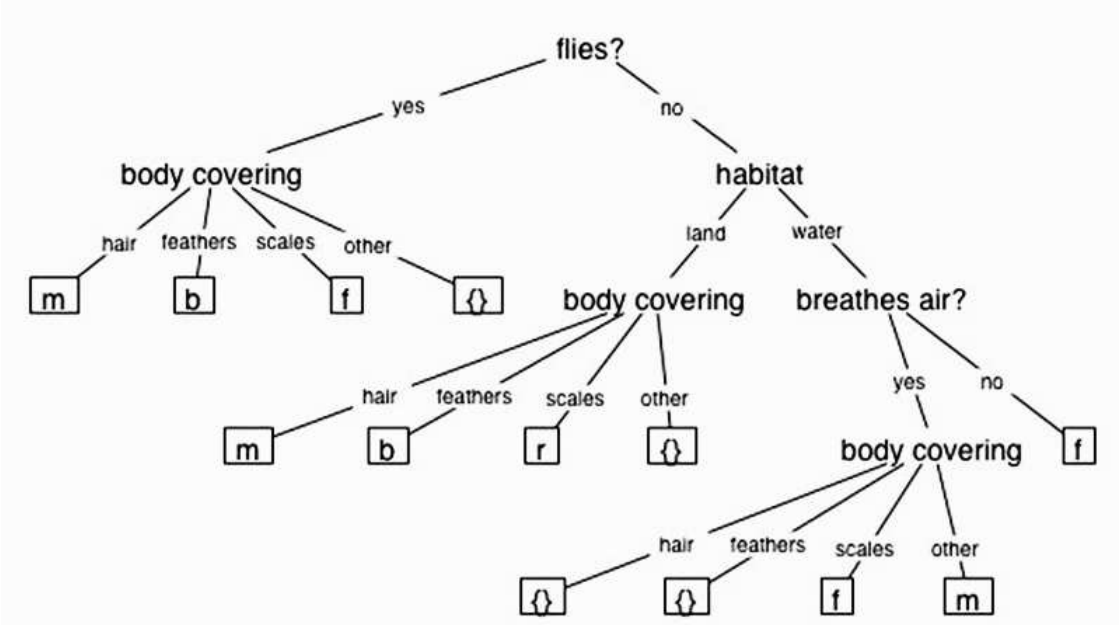


Figure 4: Example of decision tree.

In the example there are 4 attributes: flies, body covering, habitat and breathes air, and four possible categories, m, b, f and r . Here, the first attribute is flies because it is the one that produces the division on the data with minimum entropy at that level, and so on. To classify an unlabeled example you only have to follow the tree top-down, and the final leaf is the predicted category. The pathways from the root node to the leaf node can be viewed as rules, where the condition is formed by AND operation of the terms (node=arc).

C4.5 is an extension of ID3 that allows continuous numerical attributes, accounts for missing values and carries out a pruning process in order to reduce the tree size for dealing with larger amount of data. The J48 tree used in the experiments is an implementation of C4.5.

3.6. Ridor Rules Learner

Ridor stands for the RIpple-DOwn Rule learner [18]. It generates the default rule first and then the exceptions for the default rule with the least (weighted) error rate when it is used to classify the training data. Then it generates the "best" exceptions for each exception and iterates until pure. Thus it performs a tree-like expansion of exceptions and the leaf has only the default rule but no exceptions. The exceptions are a set of rules that predict the classes other than class in the default rule. IREP is used to find out the exceptions. IREP algorithm constructs rules by gradually adding one term in the condition at a time so that the error rate is minimized. The rule condition terms are like (attribute $\{=, \neq, \leq, \geq\}$ value).

4. Evaluation

The evaluation processes and measures are the same for all the experiments. Given the data, a part of the collection is considered as a training set and the remaining as a test set. So the models learn from the training set and try to predict the values of examples in the test set. Since the category of test set examples is known, we can check the predictions. Three different typical measures are calculated for each category: precision, recall and F-measure [19]. Precision is the percentage of predictions of one category that were correct. Equation 3 presents the precision expression.

$$Precision = \frac{number\ of\ correct\ predictions\ as\ category_i}{total\ number\ of\ predictions\ as\ category_i}. \tag{3}$$

Recall is the percentage of all the examples of the test set belonging to a category that were correctly predicted. The expression is presented in Equation 4.

$$Recall\ (category_i) = \frac{number\ of\ predictions\ as\ category_i}{total\ number\ of\ examples\ of\ category_i}. \tag{4}$$

F-measure is a combination of the former measures. It accounts, someway, the intersection between the examples involved in precision and recall, normalized by the sum of both. Equation 5 shows the F-measure expression.

$$Fmeasure = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}. \tag{5}$$

Thus, these three measures are calculated for each category of the test set. As said before, given the collection it is needed to divide the data in training and test sets. A common way of evaluation is cross-validation. The collection is divided into n equally sized parts. Then, each $n-1$ part combination is considered as training and the remaining part as test, so the algorithm is run $n-1$ times, and the final results are the average of this $n-1$ executions. For all the experiments described below, n has a value of 3, so training is always 66 % of the data and test stands for 33 %, running each algorithm three times. Usually, the value of n is greater than 3, typically of 10, but in this case we have very few examples of some categories, and a greater value of n could produce test sets with no representation of the mentioned categories, what is not desirable.

5. Experiments

Two kinds of experiments have been carried out. The first is intended to discover associations between attributes by considering the classification performance as an indicator of the association strength. The second kind is intended to test the prediction of future disorders.

It is needed to remark that the observations in the data set with missing values were not removed nor imputed, because the implementations of the learning algorithms are able to deal with missing data. These implementations are the ones included in the WEKA environment [20], used with default parameters to perform the experiments above.

5.1. Finding Answers

The first experiments are related to the analytical questions proposed for the Discovery Challenge of ECML/PKDD 2004 conference, specifically the ones related to the Entry collection. These tasks consist of finding relations in three different groups of patients: normal group, risk group and pathologic group. These groups correspond to the risk level of atherosclerosis - see above, and will be referenced as level groups. Specifically, the target relationships are between social factor features and physical activities features, alcohol features, smoking features, body mass index, blood pressure and HDL cholesterol, and then between physical activities and the remaining and between alcohol and the remaining. So, machine learning algorithms are applied to the data of each different group, trying to predict the value of each of the features of one group given the features of the other, viewing the possible values as the considered categories. Thus, for example, given the four social factor attributes as training factors the algorithms are run in order to predict the value of each of the four physical activities attributes and so on with the other features groups. For each relationship, the maximum values over all the different algorithms results are calculated in order to compare between level groups. So, if the prediction accuracy is good, we could say that there is a strong relationship, in a degree equal to the accuracy, between the features used for training and the feature whose values are predicted, and also we can compare prediction measures between features and level groups to state which relations are stronger than others.

Due to paper length limitations, only some of the most representative results are presented. In Figure 5, the maximum precision, recall and F-measure predictions results for ”Smoking”, given social attributes, are showed for each of the level groups, a) Normal, b) Pathologic and c) Risk, respectively, and given physical activity attributes for each of the level groups, d) Normal, e) Pathologic and f) Risk, respectively. As can be seen, in the Normal group, either from social factors or physical activity, the best prediction is reached for non-smoking people, being not significant for the remaining values of the ”Smoking” attribute. It seems that the relationship between social factors and smoking is slightly stronger than physical activity and smoking, because it produces better results for all the values of the ”Smoking” attribute. In both Pathologic and Risk groups, the relationship between the training factors and the non-smoking value is stronger for physical activity factors, being in particular high in the Pathologic group. In the latter groups, people who smoke 15 or more cigarettes a day are better predicted than in the Normal group but non-smokers are much worse detected than in the Normal group.

Let us see another representative example. Figure 6 presents the results of the prediction of cholesterol level from social factors, a), b), and c), and from physical activity factors, d), e) and f), for each of the level groups, respectively. In this case, the prediction results are very similar for the relationship between social factors and cholesterol, and physical activity and cholesterol, in all level groups, so we can conclude that the strength of the relationships is similar, too. However, it varies among level groups. In the Normal group, the mean absolute prediction error is about 24, being about 50 and 40 in Pathologic and Normal groups, respectively, concluding that it is easier to predict cholesterol, from both social factors and physical activity as training, for people in the Normal group. This fact denotes a strong relationship between the training factors and the cholesterol level in the latter group.

Finally, Figure 7 shows the results for the prediction of alcohol attribute values, separately from social factors and physical activity factors as training, for each level group, analogous to above.

The results in Figure 7 show that there is a clear relationship in all levels of groups between the training factors and the people who drink alcohol occasionally. People who drink regularly are more difficult to detect and predict from the training factors, resulting in a light relationship that is a little bit stronger in the Pathologic group. The same can be said about people who never drink alcohol in relation to physical activity factors. However, the prediction precision is significantly increased for social factors in Normal and Risk groups. People who never drink are accurately identified from their social factors in the latter groups, what denotes a significant relationship between the involved attributes.

The training features groups are taken with all the attributes at once. From the medical point of view it is also interesting to separate these attributes and to try subsets of them. So, it was tried to predict the value of the physical activity in the job attribute given all possible combinations of social factors attributes, for

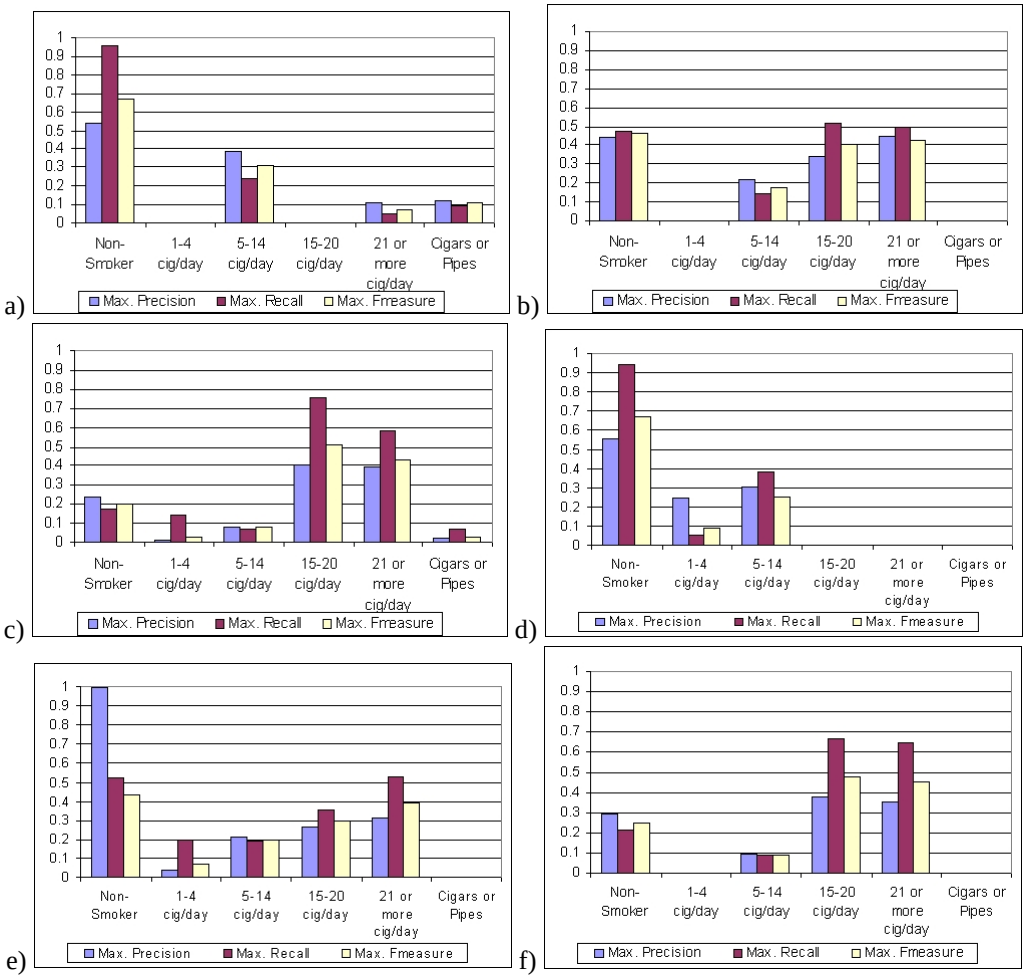


Figure 5: Maximum precision, recall and F-measure values over all the algorithms for the prediction of ”Smoking” attribute, given only the social factors on a) Normal group, b) Pathologic group and c) Risk group, and given only the physical activity factors on d) Normal group, e) Pathologic group and f) Risk group.

example. The results show that, for Normal and Risk groups, the ”Education” feature alone obtains much better prediction results than any other combination of social factors attributes. In the Pathologic group it is similar, but the difference is not so high as in the other groups, being ”Age + Education” the best combination.

5.2. Predicting Future Disorders

The main objective of the next experiments is to test the prediction accuracy of the algorithms. The Entry collection is not the only one used but also the Control collection is considered. First of all, the patients who have a control record in the Control collection, after ten years from their entry in the study, were selected. Then using their Entry attributes it was tried to predict whether they will have some disorders in ten years. These disorders correspond to systolic-diastolic hypertension, systolic hypertension, diastolic hypertension, hypercholesterolemia and hypertriglyceridemie. The possible values for these disorder attributes are true or false. The same has been done for twenty years records. The results show that the multilayer perceptron was the best algorithm, reaching values near 85 % of precision and 65 % of recall in the detection of all the disorders. The risk of future hypertension in the Risk group is 0 for many men, while some patients in this group were hypertensive since the beginning of the study. From the medical point of view it is more interesting to carry out the experiments only on the Normal group. The same process has been done separately for this group, for ten and twenty years. The results for the different mentioned disorders are

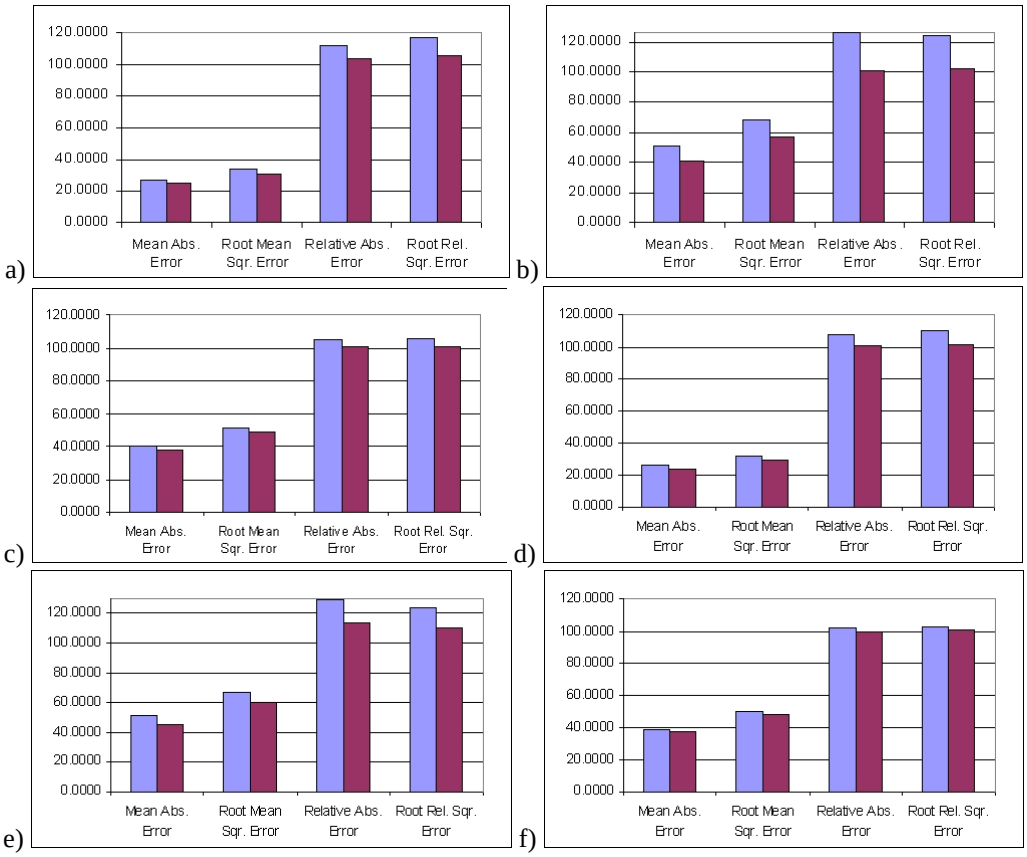


Figure 6: Average and maximum mean absolute error, root mean square error, relative absolute error and root relative square error values over all the algorithms for the prediction of cholesterol level attribute, given only social factors on a) Normal group, b) Pathologic group and c) Risk group, and given only the physical activity factors on d) Normal group, e) Pathologic group and f) Risk group.

presented in Figure 8, a) to e), for ten and twenty years prediction respectively. For each disorder, the maximum values over all the different algorithms results are presented. In this case, results show that there is not one best algorithm. Depending on the disorder to predict and also on certain categories, one algorithm fits better than others (the maximum values presented correspond to different algorithms), so it will be interesting to use all the algorithms and make decisions based on the results from all of them. As a comment, we pointed that the prediction accuracy is much higher than when entries of the three levels of groups are considered all together, confirming the early interest in the Normal group.

The values of Figure 8 show that it is more accurate to predict disorders in twenty years than to predict them in ten years, specifically the prediction of the presence of the disorders, which is accurately inferred in twenty years but very poorly predicted in ten years in all the disorders but diastolic hypertension. The non-presence of the disorders is equally well-predicted for both ten and twenty years. Among all the disorders the best detected is diastolic hypertension, obtaining prediction values near to 100 % accuracy for the presence and non-presence of the disorder. The worst predicted disorder is systolic hypertension, with the presence of the disorder non-detectable at all in ten years.

It was also tried to predict some other diseases, like angina pectoris, myocardial infarction, cerebron-vascular accident and so on, but there is a small number of observations with these features, so the results are not relevant.

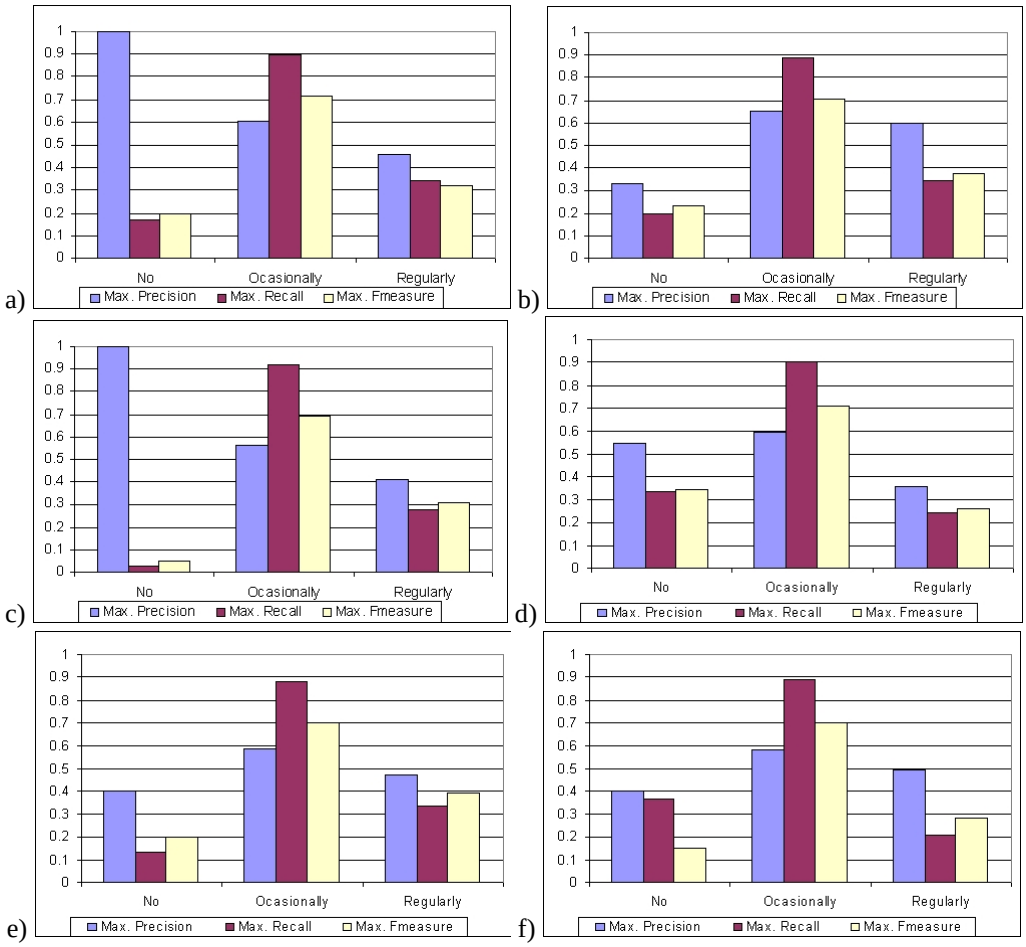


Figure 7: Maximum precision, recall and F-measure values over all the algorithms for the prediction of the ”Alcohol” attribute, given only the social factors on a) Normal group, b) Pathologic group and c) Risk group, and given only the physical activity factors on d) Normal group, e) Pathologic group and f) Risk group.

5.3. Predicting Death Cause

This experiment is analogous to the last one, but now what is tried is the prediction of the cause of death rather than diseases or disorders. Thus, the Death collection is used. The algorithms were trained with the data in the Entry collection for the patients of the Death collection. The experiments were carried out for the three levels of groups separately and for all the entries of all the groups together. The results are presented in Figure 9. In the Normal group, Figure 9 b), the best predicted causes were tumour disease and other causes. In the Risk group, Figure 9 d), the best prediction was for other causes but also for myocardial infarction and coronary heart disease, that were not predicted at all in the Normal group. In the Pathologic group, Figure 9 c), the best predicted causes were tumour disease and myocardial infarction, but stroke and general atherosclerosis could be poorly predicted, too, obtaining much lower results for these latter causes in the other groups. In general, Figure 9 a), the results of prediction of death cause are very poor, concluding that data from the Entry collection has not enough information to predict death and/or also maybe more observations are needed. But, what is sufficient for it?

6. Conclusions

Machine learning algorithms belonging to a variety of paradigms have been applied to knowledge discovery on medical data in two different ways: firstly, the methods have been used in order to predict the value of one attribute of the patient database, given a subset of other attributes as training features, proposing the

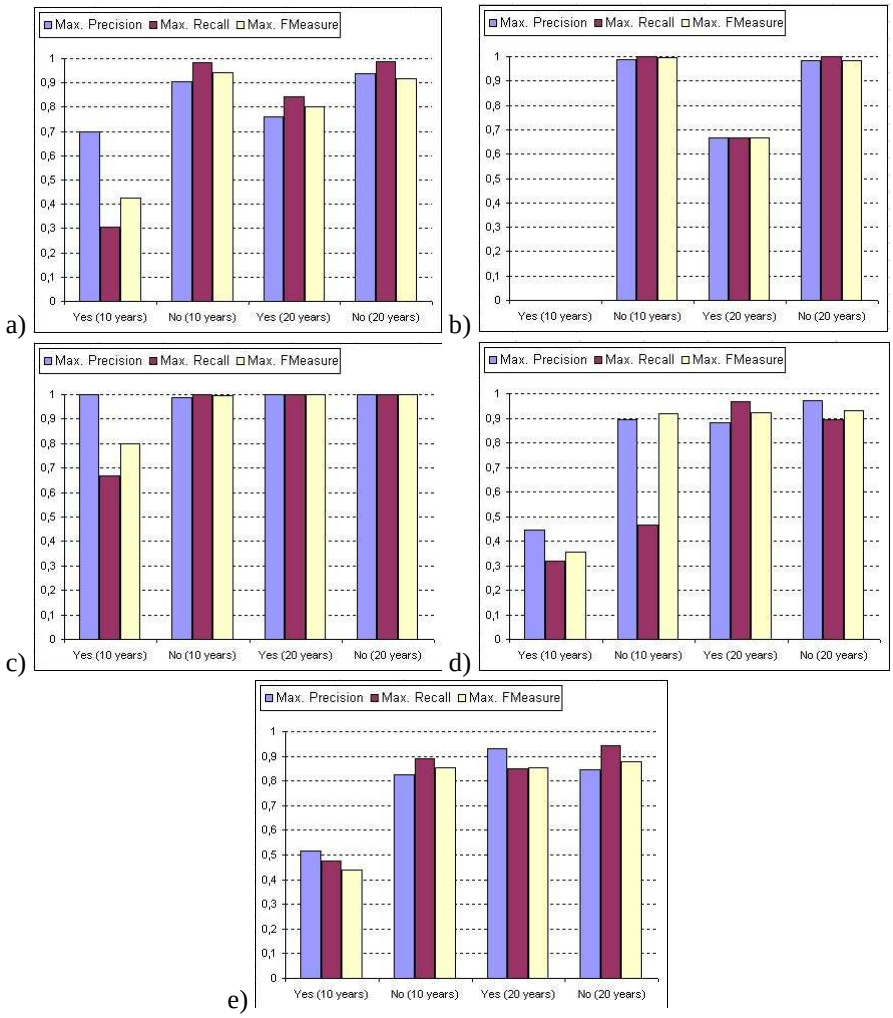


Figure 8: Maximum precison, recall and F-measure values for the 10 and 20 years predictions of a) systolic-diastolic hypertension, b) systolic hypertension, c) diastolic hypertension, d) hypercholesterolemia and e) hypertriglyceridemie.

maximum accuracy among all the algorithms as a measure of the strength of the relationship between those training features and the target attribute. This measure has been proven useful also for comparing the relationships between attributes in different groups of patients.

Secondly, the learning techniques have been applied to the prediction of future disorders. The results show that some methods predict some disorders better than others. Then, it is interesting to use all the algorithms at a time and consider the result confidence based upon the known tendency of each method. All the tested methods perform better for twenty years prediction than for ten years predictions, reaching excellent results for some of the disorders that make the methods suitable for decision support. The machine learning algorithms have been also used in the prediction of death cause, obtaining poor results in this case, maybe due to the small amount of information (entries) of this type in the dataset.

It would be interesting for the future to finely tune the parameters of the algorithms and to test more techniques. It is also intended to integrate all methods with the degree of significance and usefulness discovered in this work in order to build an expert system, and the derivation of rules understandable by humans from the results of the system will be also researched.

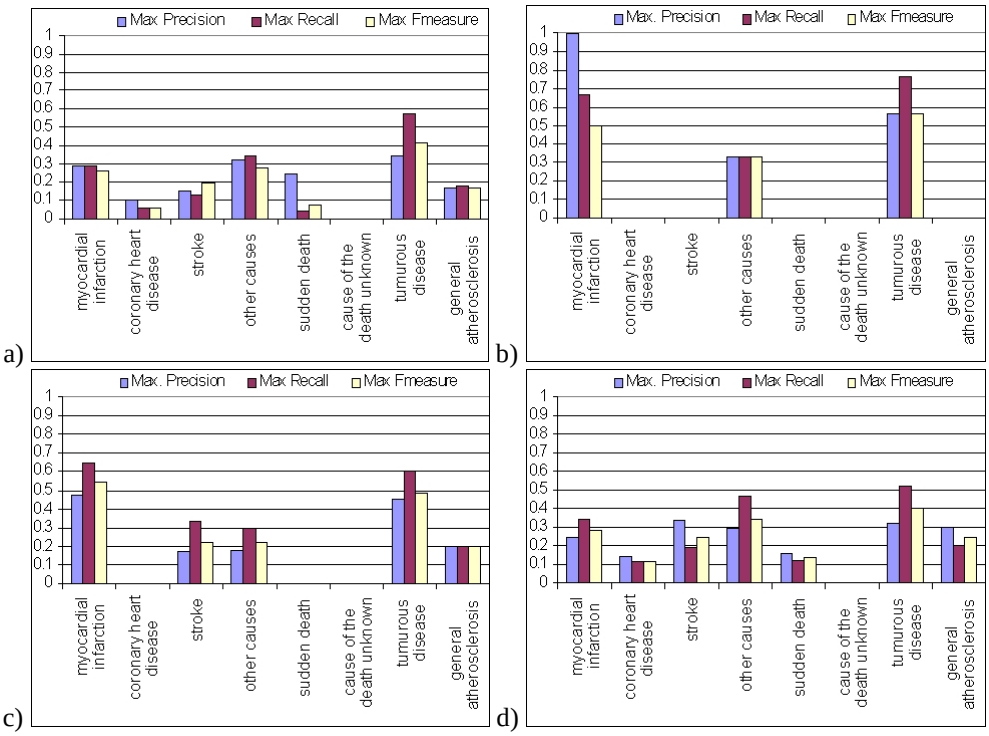


Figure 9: Maximum values of precision, recall and F-measure, in the prediction of death causes, for a) all the level groups as one, b) Normal group, c) Pathologic group and d) Risk group.

References

[1] Mitchell T., "Machine Learning", *McGraw Hill*, 1997.

[2] Lavrac N, "Selected Techniques for Data Mining in Medicine. Artificial Intelligence in Medicine", vol. 16 (1), pp. 3-23, 1999.

[3] Aseervatham S. and Osmani A., "Mining Short Sequential Patterns for Hepatitis Type Detection", *ECML/PKDD Discovery Challenge*, 2005.

[4] Aubrecht P., Kejkula M., Kremen P., Novakova L., Rauch J., Simunek M., and Stepankova O., "Mining in Hepatitis Data by LISp-Miner and SumatraTT", *ECML/PKDD Discovery Challenge*, 2005.

[5] Pizzi L.C., Ribeiro M.X., and Vieira M. T. P., "Analysis of Hepatitis Dataset using Multirelational Association Rules", *ECML/PKDD Discovery Challenge*, 2005.

[6] Durand N. and Soulet A., "Emerging Overlapping Clusters for Characterizing the Stage of Liver Fibrosis", *ECML/PKDD Discovery Challenge*, 2005.

[7] Durand N., Cleuziou G., and Soulet A., "Discovery of Overlapping Clusters to Detect Atherosclerosis Risk Factors", *ECML/PKDD Discovery Challenge*, 2004.

[8] Cios K. J., "Medical data mining and Knowledge Discovery", *Physica - Verlag*, 2001.

[9] Chen H., Fuller S. S., Friedman C., and Hersh W., "Medical Informatics: Knowledge Management and Data Mining in Biomedicine. Integrated Series in Information Systems (2)", *Springer Science and Business Media Inc.*, 2005.

[10] Boudik F., Reissigova J., Hrach K., Tomeckova M., Bultas J., Anger Z., Aschermann M., and Zvarova J., "Primary Prevention of Coronary Artery Disease Among Middle Aged Men in Prague: Twenty-year Follow-up Results", *Atherosclerosis*, 2006 Jan;184(1):86–93.

[11] Tomeckova M., "The Challenge on Atherosclerosis Data Viewed by the Experts", *ECML/PKDD Discovery Challenge*, 2004.

[12] José Ignacio Serrano, Tomeckova M., and Zvarova J., "Machine Learning Methods for Knowledge Discovery in Medical Data on Atherosclerosis", http://www.ejbi.org/, *European Journal for Biomedical Informatics*, Prague, 2006

[13] Rish I., "An Empirical Study of the Naive Bayes Classifier", *IJCAI-01 Workshop on Empirical Methods in AI*, 2001.

[14] Haykin S., "Neural Networks: A comprehensive Foundation (2nd edition)", *Pearson Education*, 1998.

[15] Scholkopf B., Smola A. J., Mtiller K. R., Burges C. J. C., and Vapnik V, "Support Vector Methods in Learning and Feature Extraction", In: Down T., Frean M., and Gallagher M., editors. Proceedings of the Ninth Australian Congress on Neural Networks, Brisbane, Australia. *University of Queensland*, 1998.

[16] Teknomo K, "K-Nearest Neighbors Tutorial", *http:people.revoledu.comkardi tutorialKNN*, 2004.

[17] Quinlan J. R., "C4.5: Programs for Machine Learning", *Morgan Kauffman*, 1993.

[18] Compton P., Edwards G., Kang B., Malor R., Menzies T., Preston P., Srinivasan A., and Sammut S., "Ripple Down Rules: Possibilities and Limitations", Boose J. H. & Gaines B. R., Ed. Proceedings of the Sixth AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop, pp. 6-1-6-20, Calgary, Canada, *University of Calgary*, 1991.

[19] Van Rijsbergen C. J, "Information Retrieval", *Butterworths, London*, 1979.

[20] Witten I. H. and Frank E, "Data Mining: Practical Machine Learning Tools and Techniques", 2nd Edition, Morgan Kaufmann, *San Francisco*, 2005.

Genetic Learning of Perceptron Networks

Post-Graduate Student:
MGR. STANISLAV SLUŠNÝ
Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic
stanislav.slusny@matfyz.cz

Supervisor:
MGR.. ROMAN NERUDA, CSC.
Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic
roman@cs.cas.cz

Field of Study:
Software systems

Abstract

This paper reviews different combinations between the most widely used type of neural networks – a multi-layer perceptron – and evolutionary algorithms. Methods to train the neural network are tested using a real-world classification problems from Proben1 benchmark suite. It is shown, that combining evolutionary algorithms with neural networks can lead to better results than relying on neural networks alone. Searching for a suitable architecture can help to find neural networks with improved performance.

1. Introduction

Artificial neural networks (ANNs) are a computational paradigm modeled on the human brain that have been applied to a variety of classification and learning tasks for a few reasons. Despite their simple structure, they provide very general computational capabilities and they can adapt themselves to different tasks, i. e. they are able to learn.

Evolutionary algorithms can be viewed as an alternative to classical optimization techniques, based on a biological metaphor: over many generations, natural populations evolve according to the principles of natural selection and "survival of the fittest", first clearly stated by Charles Darwin in The Origin of Species [1]. The basic principles of Evolutionary algorithms were first laid down rigorously by Holland [2].

In this paper, we present a comparison of different approaches to ANN learning problem. We will focus on combination of gradient and evolutionary techniques, and we will try to find optimal weights and topology of neural networks. Several experiments will be carried out on data taken from Proben1 [3] benchmark collection.

The organization of this paper is as follows. First we briefly describe perceptron networks and the core of genetic we have used. Then, the two main approaches — evolutionary learning of the weights, and evolution of network architecture — are presented. Finally, the experiments are reported and discussed.

2. Multilayer perceptron network

An ANN is an interconnected network of simple computing units called *neurons*[4]. Each neuron has n real inputs, and each input x_i has assigned a real weight w_i.

The output $y(x)$ of a neuron is defined in equation 1:

$$y(x) = g\left(\sum_{i=1}^{n} w_i x_i\right), \tag{1}$$

where x is the neuron with n input dendrites (x_0 ... x_n), one output axon $y(x)$, (w_0 ... w_n) are weights and $g : \Re \rightarrow \Re$ is the activation function that determines how powerful the output (if any) should be from the neuron, based on the weight sum of the input.

One of the most common activation function is a logistic sigmoid function 2

$$\sigma(\xi) = 1/(1 + e^{-\xi t}), \tag{2}$$

where t determines its steepness. Very important property of the sigmoid function is smoothness and the fact, that it has an easily calculated derivative.

In a *multilayer feedforward* ANN, the neurons are ordered in layers, starting with an *input layer* and ending with an *output layer*. Between these two layers there can be a number of *hidden layers*. Connections in this kind of networks only go forward from one layer to the next. Neurons in input layer are called *input neurons*, similarly neurons in output layer are called *output neurons*. Let us denote n_I, n_O, n_H number of input, output and all hidden neurons, respectively.

Multilayer feedforward ANNs work in two different phases: In a *training phase* (learning phase) the ANN is trained to return a specific output given a specific input. In the *execution phase* the input is presented to the input layer, propagated through all the layers (using equation 1) until it reaches the output layer, where the output is returned. The *architecture* of ANN defines the number of layers, number of neurons in each layer and connections between neurons.

We will focus on a learning situation known as a *supervised learning*, in which a set of input/desired-output patterns is available. The training process can be seen as an optimization problem, where we wish to minimize the *mean square error* (E_{MSE}) of the entire set of training data. E_{MSE} is defined as the squared error of the ANN for a set of patterns:

$$E_{\text{MSE}} = \sum_{k=1}^{p} E_k(w) \tag{3}$$

$$E_k(w) = \frac{1}{2} \sum_{j \in Y} (y_j(w, x_k) - d_{kj})^2, \tag{4}$$

where Y is a set of output neurons, p is the number of patterns in the set, y_j is an output of neuron j given weight vector w and k-th input pattern x_k, and d_{k_j} is desired output of output neuron j for input pattern k.

The *classification problem* is that of assigning an input vector to one of M classes. Each pattern from the training pattern set contains an input vector and its desired output vector. The output of the network must be interpreted as a class. Such interpretation can be performed in different ways. One of them consists in assigning an output neuron to each class — when an input vector is presented to the network, the network response is the class associated with the output neuron with the largest output value. This method is known as the *winner-takes-all* approach.

For classification problems, E_{CLAS} reports — in a high-level manner — the quality of the trained ANN and is defined as a percentage of incorrectly classified patterns.

3. Evolutionary algorithm

Evolutionary algorithm has been investigated by John Holland [2], with a marked increase of interest within the last few years. Evolutionary search refers to a class of stochastic optimization techniques — loosely

based on processes believed to operate in biological evolution — in which a population of candidate solutions (chromosomes) evolves under selection and random "genetic" diversification operators. The problem is to find minimum or maximum of *fitness function*. Every member of population is called *individual* and represents a potential solution to a problem.

The algorithm 1 reveals skeleton of genetic algorithm used in our experiments. To fully describe genetic algorithm, it is needed to define how each solution will be represented, how initial population will be created and what genetic operators will be used in the Reproduction step.

Input: number of individuals in population N, number of elits E, maximum number of populations G_{max}.
Output: the best found solution of a problem.

1. Start: Create initial population of N individuals $P(0) = \{I_1, ..., I_N\}$, $i = 0$.
2. Evaluation of individuals: To compute fitness function for every individual I, build ANN corresponding to I and compute E_{MSE} for a training set. Let $\mathcal{F}(I) = E_{\text{MSE}}$.
3. Stop-condition: If $i = G_{max}$, finish and return solution represented by individual with minimal value of fitness function.
   (a) Creation of new population $P(i+1)$ from population $P(i)$: Create empty population $P(i+1)$.
   (b) Selection: Choose E best individuals from population $P(i)$ and move them to population $P(i+1)$. With chosen operator of selection choose $N - E$ individuals and insert them to population $P'(i)$.
   (c) Reproduction: Apply chosen operators on population $P'(i)$, the result is population $P(i+1)$.
      i. Application of binary operators: If population $P'(i)$ contains odd number of individuals, insert chromosome of the first individual from population $P'(i)$ into population $P''(i)$. From population $P'(i)$ choose pairs of chromosomes C and D and apply on them reproduction operators, new chromosomes C' and D' insert to population $P''(i)$.
      ii. Application of unary operators: On every chromosome from population $P''(i)$ apply chosen unary operator, new chromosome insert to population $P(i+1)$.
4. New generation: $i = i + 1$, Continue by step 2.

Table 1: Skeleton of the genetic algorithm.

We have used the *roulette wheel selection* in all experiments. The selection operator performs the equivalent role to natural selection — it chooses individuals for next population proportionally to their fitness values. As we wanted to minimize the error function in all our experiments, the probability p_i of selection i-th individual is defined by the equation 5

$$p_i = \frac{(1+\varepsilon) * C - \mathcal{F}(I_i)}{N * (1+\varepsilon) * C - \sum_{j=1}^{N} \mathcal{F}(I_j)}, \tag{5}$$

where N is the number of individuals and C is the maximal fitness value in the population, and ε is a small positive constant, which ensures that probability of selection the worst individual will be non-zero.

4. Evolution of weights

Weight training in ANN is usually formulated as a minimization of error function, and is carried out by some gradient descent algorithm such as Back-Propagation, or one of its many variants[5]. Due to its use of gradient descent, these algorithms often get trapped in the local minimum of the error function,

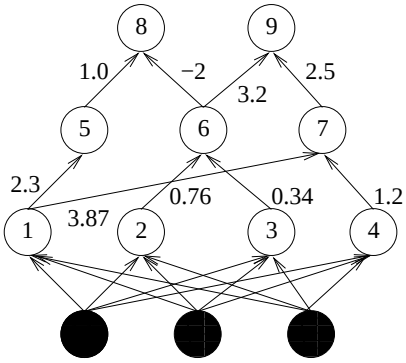


Figure 1: Example of ANN with seven hidden units in two hidden layers, three input and two output neurons.

$$\begin{bmatrix} \otimes & \otimes & \otimes & \otimes & 2.3 & \otimes & 3.87 & \otimes & \otimes \\ \otimes & \otimes & \otimes & \otimes & \otimes & 0.76 & \otimes & \otimes & \otimes \\ \otimes & \otimes & \otimes & \otimes & \otimes & 0.34 & \otimes & \otimes & \otimes \\ \otimes & \otimes & \otimes & \otimes & \otimes & \otimes & 1.2 & \otimes & \otimes \\ \otimes & \otimes & \otimes & \otimes & \otimes & \otimes & \otimes & 1.0 & \otimes \\ \otimes & \otimes & \otimes & \otimes & \otimes & \otimes & \otimes & -2 & 3.2 \\ \otimes & \otimes & \otimes & \otimes & \otimes & \otimes & \otimes & \otimes & 2.5 \end{bmatrix}$$

Figure 2: Example of encoded ANN from the previous figure.

and are incapable of finding global minimum. As the evolutionary algorithm need not to use the gradient information and works with population of potential solutions, it has a better ability to avoid local minimum trap. What more, evolutionary algorithm can be applied in situations, when gradient information is not available at all, it can handle global search problem in a vast, complex and non-differentiable surface. The evolutionary approach also makes it easier to generate ANNs with some special characteristics. For example, ANNs complexity can be decreased and its generalization ability increased by including a special term in the fitness function penalizing large networks.

4.1. Representation

The standard genetic algorithm uses binary strings to encode alternative solutions. In a such representation scheme, each connection weight is represented by a number of bits of certain length. The advantages of this representation are mainly simplicity and generality. It is possible to use well known crossover (such as one-point crossover) or mutation operators for binary strings. Although there are several encoding methods, that encode real numbers with different range and precision, trade off between precision and range often has to be made. Real-world experiments demand big precision, what causes too long chromosome for which the evolution process becomes non-efficient. Therefore, different encoding method is used. The chromosome is interpreted as a matrix of dimensions $n_H \times n$, where $n = n_O + n_H$. In the i-th row and the j-th column there is either a special symbol $\otimes$, if neurons i and j are not connected, or the weight of connection from i to j. The following connection matrix from figure 2 encodes the ANN from the figure 1. The chromosome is then created by concatenating all the rows from the matrix. In case of feedforward ANNs, the entry on i-th row and j-th column can be non-zero only for $i < j$. This reduces the length of the chromosome and the evolutionary process becomes more effective.

4.2. Initial population and operators

Initial population consists of fully connected neural networks with randomly initialized values. To each connection a weight from $[-1.0, 1.0]$ interval is assigned.

For training the weights of ANN, we experimented with operator *Biased-Weights-Mutation*[6], which adds

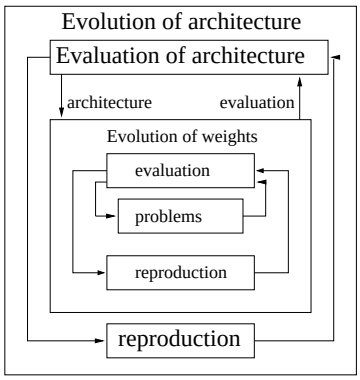


Figure 3: Schema of evolutionary algorithm searching optimal architecture of ANN.

to the weights of randomly selected connections in the parent network with values chosen randomly from the same probability distribution used for initialization. *Weights-Crossover* performs uniform crossover with the basic units exchanged being the set of all incoming weights of a particular node.

5. Evolution of architecture

In the previous section, we have assumed that the evolved ANN has a fixed architecture. Selection of a good architecture is state of art — although it has significant impact on network's information processing capabilities, there are not known satisfiable rules, on how to choose a good one for a particular task. ANN with only a few connections may not be able to satisfiable learn the task because of its limited capacity, while ANN with large number of connections and neurons may over fit and fail to have a good generalization ability.

Similarly to previous section, the problem of determining optimal architecture for a particular task can be formulated as an optimization problem. As stated in [7], evolutionary algorithms are a good method for searching in such a complex surfaces. In the case of *simultaneous evolution of weights and architecture* both weights and characteristics of architecture are encoded in the chromosome. As shown on figure 3, this method contains the previously solved problem of determining optimal set of weights as a subtask. On the other side, the problem of *separated evolution of weights and architectures* is the noisy fitness evaluation, caused by random initialization of weights. Only architecture is encoded in chromosome, weights have to be learned later. The fact that an individual gained better fitness value does not mean that this individual is really better — in is necessary to repeat the evaluation and average the obtained values. This way, the computation time increases dramatically.

5.1. Representation

We used the *indirect encoding scheme* in our experiments [7]. Chromosome consists of several sections, which hold information about some important characteristic of ANN's architecture. In the first section is stored the number of hidden layers, follows section with information about the number of neurons in each hidden layer and in the case of simultaneous evolution of weights and architecture, the section with connection matrix is appended. The operators used in evolutionary algorithm must be aware of this representation — operators from previous section are allowed to modify only the last section of the chromosome. To modify the remaining parts, new operators are introduced.

5.2. Operators

There are two operators that work on the level of units. The *Remove-Neuron* operator removes randomly selected neurons. Creation of new neurons is done by the *Duplicate-Neuron* operator, which selects random neuron and duplicates it. Similarly, there are two operators working on connections, the *Remove-*

Number of hidden layers.
Number of neurons in the first hidden layer. ... Number of neurons in the last hidden layer.
Connection matrix.

Table 2: Chromosome for evolving architectures of ANN.

Connection operator removes random number of connections from each neuron,while the *Add-Connection* operator randomly adds connection between two neurons.

6. Experiments

A set of experiments has been carried out on several data sets from the Proben1 [3, 8, 9] benchmark. This way we are able not only to provide relative comparison, but also to explore the efficiency of the algorithms with respect to the best results obtained by other methods and authors. In the following we briefly describe several experiments and try to generalize the results.

We have chosen four classification problems: The goal of ANN is to classify tumor as either benign or malignant in the Cancer problem, diagnose diabetes of Pima indians in Diabetes problem, predict the heart disease in Heart problem and recognize one of 19 different diseases of soybean in Soybean problem. Characteristics of each data set are shown in table 3.

	Classes	Examples	b	c	n	Tot.	b	c	m	Tot.
Cancer	0	9	0	9	0	9	0	9	2	699
Diabetes	0	8	0	8	0	8	0	8	2	768
Heart	1	6	6	13	18	6	11	35	2	920
Soybean	16	6	13	35	46	9	27	82	19	683

Table 3: Problems and the number of binary, continuous, and nominal attributes in the original dataset, number of binary and continuous network inputs, number of network inputs used to represent missing values, number of classes, number of examples.

The results of experiments are reported in terms of the values of the two errors, E_{MSE} and E_{CLAS} measured both on the training set and previously unseen test set. In the following tables we use symbols M_t (or M_s) for E_{MSE} over training (or testing) set, and C_t (and C_s) for E_{CLAS} over training (testing, respectively) set.

6.1. Searching for suitable connections

This experiment tested the separated evolution of architecture and weights. Thus, the algorithm consists of two steps - in the first one the direct architecture encoding described in Tab. 2 is used and evolved. To determine a fitness of such an individual, several (30 in our case) randomly initiated runs of rprop algorithm are carried. In the second step, the evolved architecture is trained by full-fledged run of a back propagation algorithm (10 times for different random weights initializations). Results can be compared to Tab. 4, which gathers results of the classical back propagation, the rprop gradient learning algorithm and GA using the Biased-Weights-Mutation and Weights-Crossover (cf. Tab. 4). Typically we ran the gradient algorithm for 5000-25000 epochs, while the GA ran for 1000-2000 generations (with population of 300 individuals).

Results of this experiment were quite satisfiable (cf. Tab. 5): GA was able to find better architectures, containing less connections, and achieving better training error. As a side effect, these specialized architectures usually achieved worse generalization error, in comparison to fully connected architectures, which is understandable (compare Tab. 4 and Tab. 5).

	Diabetes			Cancer		
	BP	RP	GA	BP	RP	GA
M_t	0.2675	0.2473	0.2882	0.0463	0.0390	0.0484
C_t	0.2645	0.2616	0.2083	0.0280	0.0438	0.0286
M_s	0.3217	0.3154	0.3268	0.0242	0.0180	0.0159
C_s	0.2645	0.2615	0.2266	0.0280	0.0438	0.0057
	Soybean			Heart		
	BP	RP	GA	BP	RP	GA
M_t	0.0150	0.2112	0.7416	0.1840	0.0904	0.2372
C_t	0.2536	0.3080	0.5585	0.2033	0.2079	0.1536
M_s	0.1196	0.2907	0.7474	0.2574	0.2559	0.2904
C_s	0.2536	0.3078	0.5513	0.2033	0.2077	0.1913

Table 4: Error comparison for back propagation (BP), rprop (RP) and genetic algorithm (GA) used for weights learning.

	Diabetes	Cancer	Soybean	Heart
M_t	0.4541	0.0436	0.0069	0.0566
C_t	0.3516	0.0157	0.1615	0.1764
M_s	0.4568	0.0238	0.1369	0.2501
C_s	0.3516	0.0157	0.1615	0.1764

Table 5: Search for connections: Error results for architectures recommended by GA.

6.2. Searching for a suitable architecture

The goal of the second experiment was to combine the search for architecture with the evolution of weights in one GA. We have started with individuals with random number of hidden layers (between 1–4), which remained constant during the algorithm. However, the number of neurons and connections were varying, as well as the values of weights. The following genetic operators we used:

operator	p_m
Duplicate-Neuron	0.05
Remove-Neuron	0.05
Add-Connection	0.05
Remove-Connection	0.05
Biased-Weights-Mutation	0.1

Table 6: Operators probability.

The overall results were very satisfiable (cf. Tab. 7, compare with Tab. 4). The GA was able to evolve suitable architectures for a given task. It is interesting, that most of the solutions had just one hidden layer, some of them had two. In some papers authors use penalization for more complex solutions, which was not necessary here, because the evolution tends to exclude more complex networks based on their fitness anyways. The architectures recommended by the GA provided even better results than the ones reported as best (found by human) in [3].

7. Conclusions

As shown, evolution can be introduced into ANN learning problem at different levels. Suggested algorithms were tested on real-world problems from Proben1 benchmark suite. The evolutionary algorithm is a complex and robust method, which can be used to search both optimal weights and architecture of ANN. Although the evolutionary process can be easily parallelized, computation is always very time consuming. The results obtained by simultaneous evolution of weights and architecture were surprisingly good. Not

	Diabetes	Cancer	Soybean	Heart
M_t	0.2679	0.0382	0.0176	0.1025
C_t	0.1927	0.0190	0.0088	0.0551
M_s	0.3108	0.0184	0.0989	0.2342
C_s	0.2266	0.0057	0.0616	0.1435

Table 7: Search for an architecture and weights: Error results for genetically evolved network architectures.

only the resulting ANNs gained excellent results on the training set, they showed a good generalization ability. The complexity of found architecture for a particular task mirrored it's real difficulty.

The combination of evolutionary techniques and ANNs can lead to better intelligent systems, than relying on ANNs alone. With the increasing power of parallel computers, the evolution of large ANNs becomes feasible.

References

[1] Ch. Darwin. The Origin of Species, New American Library, Mentor paperback, 1859.

[2] J. Holland. Adaptation In Natural and Artificial Systems. The University of Michigan Press, Ann Arbour, 1975.

[3] L. Prechelt. PROBEN1 - A set of benchmarks and benchmarking rules for neural network training algorithms. Technical Report 21/94, Universitat Karlsruhe, Germany, September 1994.

[4] J. Šíma, R. Neruda. Teoretické otázky neuronových sítí, Matfyzpress,

[5] M. Riedmiller, H. Braun. A direct adaptive method for faster backpropagation learning: The RPROP algorithm, Proceedings International Conference

[6] D. Montana, L. Davis. Training feedforward neural networks using genetic algorithms, in Proc. 11th Int. Joint Conf. Artificial Intelligence. San Mateo, CA: Morgan Kaufmann, 1989, pp. 762-767. P

[7] X. Yao, Y. Liu. Evolving neural network ensembles by minimization of mutual information,' International Journal of Hybrid Intelligent Systems, 1(1):12-21, January 2004. raha, 1996.

[8] O. L. Mangasarian, W. H. Wolberg. Cancer diagnosis via linear programming, SIAM News, Volume 23, Number 5, September 1990, pp 1 & 18.

[9] R.S. Michalski, R.L. Chilausky. Learning by Being Told and Learning from Examples: An Experimental Comparison of the Two Methods of Knowledge Acquisition in the Context of Developing an Expert System for Soybean Disease Diagnosis, International Journal of Policy Analysis and Information Systems, Vol. 4, No. 2, 1980.

[10] J. R. Koza, J. P. Rice. Genetic Generation of Both the Weights and Architecture for a Neural Network, Proceedings of the IEEE Joint Conference on Neural Networks, pp. 397-404, 1991.

[11] H. Kitano. Designing neural networks using genetic algorithms, 1989.

[12] C. Forgaty, T., R. Huang. Implementing the genetic algorithm on transputer based parallel processing systems. In Proceedings of Parallel Problem from Nature, pages 145-149, 1991.

Security Model Based on Virtual Organizations for Distributed Environments

Post-Graduate Student:
ING. ROMAN ŠPÁNEK

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8
Czech Republic

Department of Software Engineering
Faculty of Mechatronics and Interdisciplinary Engineering Studies
Technical University of Liberec
Hálkova 6
161 17 Liberec 1

Czech Republic

spanek@cs.cas.cz
roman.spanek@vslib.cz

Supervisor:
ING. JÚLIUS ŠTULLER, CSC.

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8

Czech Republic

stuller@cs.cas.cz

Field of Study:
Electrotechnics and Informatics

The work was supported by the project 1ET100300419 of the Program Information Society (of the Thematic Program II of the National Research Program of the Czech Republic) “Intelligent Models, Algorithms, Methods and Tools for the Semantic Web Realization”, partly by the Institutional Research Plan AV0Z10300504 “Computer Science for the Information Society: Models, Algorithms, Applications” and by the Ministry of Education, Youth and Sports of the Czech Republic, Project No. 1M4674788502 “Advanced Remedial Technology and Processes”.

Abstract

The paper presents a new approach for treating security issues in various environments with special accent on Mobile databases, Semantic web and Grids. Target environments’ specifications imply that a security model should have at least low space complexity, some level of autonomy and should be implementable in a distributed manner. To achieve all these requirements we propose a security model based on virtual organization that has ability to preserve and manage the trust between users. The model is build up on mathematical background based on hypergraphs that are the way how to reduce space complexity of the model and how to improve and simplify the implementation. To verify proposed model an experimental implementation SecGRID was programed and some results are presented.

1. Introduction

Rapid evolution in many computing areas brings up many useful aspects, but also many problems and issues to be addressed. In this paper we will concentrate on tree distributed environments namely:

- Mobile databases
- Semantic web
- Grid computing

Although all three have some different features, they also have a lot of in common. Before we proceed to these common features, let us briefly overview all three environments.

Mobile databases [1], [2], [3], offer the ability to access and exchange information anywhere, at any time. The possible network architectures can be summarized as:

- cellular networks
- multihop wireless networks - broadly known as ad hoc networks
- sensor networks

In the first case of *cellular networks*, some specialized nodes, called base stations, coordinate and control all transmissions within their coverage area - *a cell*. The base station grants access to the wireless channels in response to service requests received from the mobile nodes currently in its cell.

The primary characteristic of an *ad hoc network* architecture [4], [5], [6], [7] is the absence of any stationary structure. Ad hoc nodes can communicate directly with the nodes in their transmission range in a peer-to-peer fashion. Communication to distant nodes is achieved in multi-hop fashion. Therefore each ad hoc node acts also as a router, storing and forwarding packets on behalf of other nodes. The result is a generalized wireless network that can be rapidly deployed and dynamically reconfigured to provide on-demand networking solutions.
Besides the fact that mobile databases have a lot of advantages, they also have some challenges to cope with. Taking issues like power supply limits, limited bandwidth and unreliability of wireless lines, the security is one of the most important. Were an efficient and strong security solution not to be available, it would be very hard to achieve all advantages of ad hoc networks.

Sensor network can be characterized as networks build up from tiny sensing units having some communication and computation capabilities. The whole sensor network lifetime comprises three main phases. During the first, *deployment phase*, sensors are deployed in the sensor field. In the second phase, *post-deployment phase*, sensors have to build up a topology and then they have to cope with potential changes in it. This phase is in particular important as the topology changes continually for available energy, malfunctioning, etc. During the last phase, *redeployment phase*, some additional sensors may be deployed to replace malfunctioning nodes. Nodes in sensor network sense their vicinity for desirable phenomena and in a positive case sensors cooperate in multi-hop communication fashion to deliver data to a unit responsible for its further processing. As sensor networks are a bit specific, we do not addressed them in the paper anymore.

The *Semantic Web* is often believed to be the successor of the current web. Its main idea is to describe resources in the form of machine processable meta-data allowing automation of the requested tasks connected with the retrieval and usage of these resources. Although the main focus of previous work was aimed at the creation of knowledge representation languages (RDF-S, DAML+OIL [8], OWL [9]), reasoning systems, and also at the tools helping to embed web pages with semantic markup, the emerging commercial applications such as e-commerce, banking or travel services face a lot of security issues. Without a secure solution, it would be very hard to exploit all promising features of semantic web vision. The first possible approach is to extend the current security mechanisms used in distributed systems (Kerberos [10], PGP [11], SPKI [12] etc.). These technologies, however, cannot be seamlessly transferred due to the fully decentralized nature of the web, extremely large number of resources, services, agents and users, and their heterogeneity. Moreover, the number of entities accessing resources and interacting with themselves can be very large and can rapidly change.

The *Grid computing* paradigm can be characterized by a large number of interconnected users and sites cooperating on a common task. Users in a Grid are usually organized in *Virtual Organizations* (VOs). A Virtual Organization is a temporary or permanent coalition of geographically dispersed individuals, groups, organizational units or entire organizations that pool resources, services and information to achieve common objectives. The Dynamic Virtual Organizations Membership and structure of such a VO may evolve over time to accommodate changes in requirements or to adapt to new opportunities in the business environment. Considering this, it is straightforward that grid computing strategies can be used in the web environment for security improvements.

Even thought the mentioned areas do have some specifications typical for them, such as huge amount of pages in the case of semantic web, mobility of users in mobile computing paradigm, or heterogeneity of connected sources in grids, they also have some common specifications. Further, while all of them offer ability to share resource, support communication and cooperation between users (but not only the human users), the security is the crucial issue being common for all mentioned areas. Therefore it is natural to expect solutions that try to solve the problem with the security for all of them.

The rest of the paper is organized as follows: section 2 briefly introduces the related security models. Our security model is then described more in details in the next section. Section 4 contains some experimental results. The paper is then concluded.

2. Security Models and Approaches

The security is one of the key concerns in many areas. The security level is usually one of the key factors influencing the usability of any proposal not only in computer science. Therefore, it is natural that security has been given attention by researches and many solutions and proposals for improving the security have been introduced. In general, two main levels of the security are:

- cryptography level
- trust level

It is important to note that these levels are not isolated one from another, but they are very often used complementary. A brief description of both levels follows.

2.1. Cryptography

On the first level strong cryptography algorithms are the basis, taking responsibility for shielding transmitted data against man-in-the-middle attack, threat of tap, etc. Cryptography plays also important role in certificates (Public Key Infrastructure - PKI [13]) enabling users to communicate, share resources and information, etc. Cryptography, however, is enough only when considering tasks like sending messages, sharing files, etc. with accent on secure transmission of sent data. This approach, nevertheless, suffers by lack of additional abilities required by human users, like when to share data, when to trust the sender, etc. Therefore the next level is responsible for the trust management.

2.2. Trust Models

The trust management approaches build an enhanced security level on underlying cryptography level. The main task is to build, preserve and manage relationships between users. The relationship are usually build up on the trust.

The trust of a party A to a party B for a service X is a measurable belief of A in that B behaves dependably for a specified period within a specified context (in relation to service X)

This approach is similar to the well known term of creating Virtual Organization in the grid environment. The necessary condition for practical evolution of VOs is to have a strong mechanism preserving their overall security.

2.3. Related Work

In the following a brief overview of the trust management approaches proposed for VO is given. Two main approaches are currently available for the *trust management*:

The policy-based approach has been proposed in the context of open and distributed services architectures [14], [15], [16], [17], [18] as well as in the context of Grids [19] as a solution to the problem of authorization and access control in open systems. Its focus is on trust management mechanisms employing different

policy languages and engines for specifying and reasoning on rules for trust establishment. In addition, it is possible to formalize trust and risk within rule-based policy languages in terms of logical formulae that may occur in rule bodies. Currently, policy-based trust is typically involved in access control decisions. Declarative policies are very well suited for specifying access control conditions that are eventually meant to yield a boolean decision (the requested resource is either granted or denied). Systems enforcing policy based trust typically use languages with well-defined semantics and make decisions based on "nonsubjective" attributes (e.g., requester's age or address) which might be certified by certification authorities (e.g., via digital credentials). In general, policy-based trust is intended for systems with strong protection requirements, for systems whose behavior is guided by complex rules and/or must be easily changeable, as well as for systems where the nature of the information used in the authorization process is exact.

The reputation-based approach has emerged in the context of electronic commerce systems, e.g. eBay. In distributed settings, reputation-based approaches have been proposed for managing trust in public key certificates, in P2P systems, mobile ad-hoc networks, and recently, in the Semantic Web, such as [20], [21], [22], [23], [24], [25]. Typically, the reputation-based trust is used in distributed networks where any involved entity has only a limited knowledge about the whole network. In this approach, the reputation is based on recommendations and experiences of other users/sites.

In the following we will put a strong emphasis on creating the underlying VO by "evolution". In order to describe it we need a model which can efficiently capture VO changes. This model is described in the subsequent section.

3. Security Model Based on Virtual Organization

As was mentioned previously, VO can be useful model for treating the trust between users in all related environments (mobile databases, semantic web, grids). On the other hand, VO model can be limited by some specific features of the environment:

- *mobile database* environment in addition to mobility of users also poses severe limitations to storage and computation capabilities of devices
- *semantic web* environment with almost unlimited number of users poses requirements on storage complexity
- *grids* with heterogeneity and geographic diversity of resources

this list of additional limitations showns that a model having the following specification is required:

1. The model should be able to store large amount of users in low storage complexity. VOs are very often modeled and depicted as (oriented) weighted graphs. But the complexity of storing information about all members in VO might be very high. The given complexity is $O(n^2)$, where n is the amount of vertexes. This is, however, unacceptable in case of mobile database environment and also semantic web might very quickly exceed storage capacity of particular node.
2. The model should have some level of autonomy in building relationships and the trust among users. The autonomous feature of the model is crucial when considering environments where users' relationships became complicated or agent technologies are used. Such feature is highly useful when users would like to create strongly connected groups *on-the-fly*. Nowadays approaches usually assume that such groups are created by somebody and usually manually. We consider such creation as a bottleneck of these models.
3. The model should be implementable in a distributed (heterogeneous) environment. A distributed implementation is the key factor influencing model capabilities and usefulness.

Our approach is therefore build up on the previous list of requirements. The next section describes the very base of our proposal.

3.1. The Security Model

We believe that the security model that fulfill all listed requirements from the previous section should have an appropriate mathematical background. The reason is, that the background can simplify the implementation and also improve efficiency by offering a set of well known techniques for storing, organizing and optimizing implementations. As a consequence, let us shortly describe the used mathematical model of hypergraphs.

Hypergraph is quadruple (V, E, W_v, W_e), where V is a set of vertexes, E is set of edges ($E \in 2^V$), W_v is a set of vertexes' weights and finally W_e is a set of edges' weights. The main difference between graphs and hypergrahs is that an hyperedge can be incident to more then two vertexes. A hypergraph is in Figure 1. It is example of hyperpraph containing 5 hyperedges and 8 vertexes.

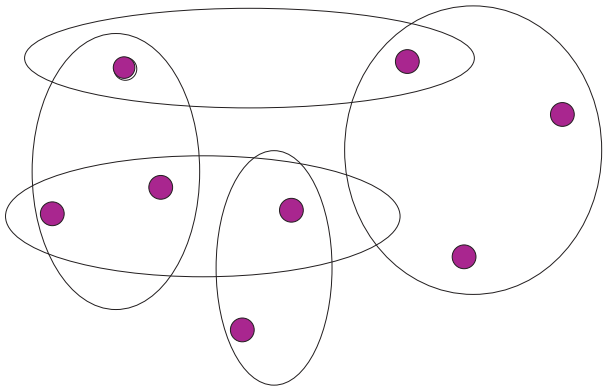


Figure 1: An example of hypergraph.

In Figure 2 the same situation is sketched, but now using graphs instead of hypergraphs. The edges in Figure 2 are shown in different colors and styles according to hyperedges from Figure 1.

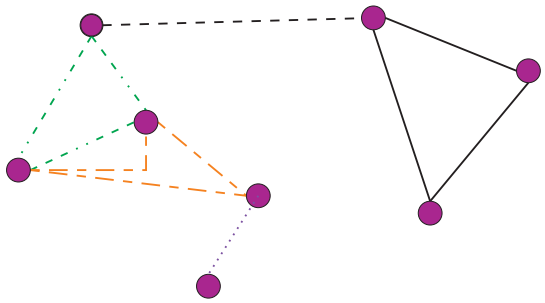


Figure 2: An example of graph showing groups of vertexes.

It is clear that hypergraphs are generalization of graphs. It is also clear that hypergraphs are richer structures than graphs. On the other hand, the richness of hypergraph brings some implementations issues (but these issues are out of scope of the paper).

Let now return to the list of three additional limitations from the previous section.

- The first item on the list was the low space complexity. In the case of graphs the space complexity is $O(n^2)$, which is unacceptable. On the other hand, in Figure 1 one can see how VO can be stored as hypergraph. In such a case VO is not stored as sets of vertexes, edges and their incidences, but simply by a membership of hyperedges. Therefore hypergraphs can be very useful for modeling VO reducing the space complexity.
- The second requirement is a kind of autonomy. One of possible solutions is to have set of rules that take care of all edges and also vertexes in the VO. Hence, we proposed a set of rules for reevaluating edges' and vertexes' weights (see [26].
- Third item on the list requires implementation in distributed environment. When trying to build up a list of all possible distributed implementation, one should start with implementations based on *Remote Procedure Call* (RPC), like CORBA[27] or JavaRMI [28]. Another technology worth mentioning are *services*. As an example let us mention web services based on WSDL[29] and SOAP[30]. One of the last possibility is to use *message passing*. The main advantage of message passing is simple and environmentally independent implementation. With respect to our needs and also to target environments the best choice is message passing with its simple, straightforward and efficient implementation.

4. Experimental Application SecGRID

An experimental implementation SecGRID was programed in ANSI/C and its aim was to verify that proposed algorithms for edge reevaluation preserve consistency of the VO. By the consistency of the VO we mean that the structure will:

- *not degenerate to one huge VO containing all nodes*
- *not degenerate to huge amount of very small VOs*
- *preserve relationships (expressed by an edge weight) between users*

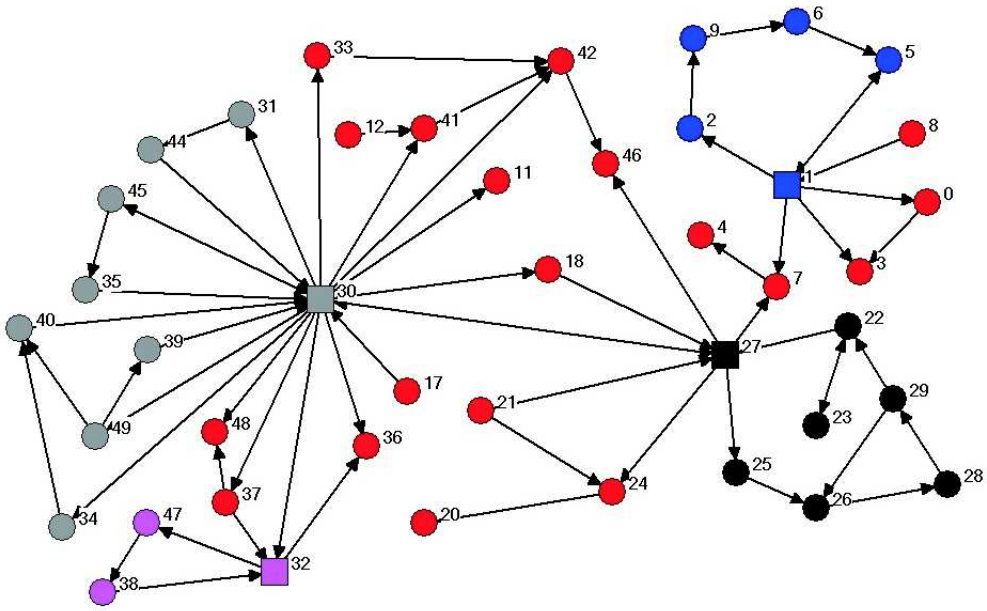


Figure 3: Experimental implementation output.

To verify the model the simulation comprised two phases:

1. random generation of edges
2. intentional edge generation with special accent on problematic situation in the VO evolution

One of the resulted graphical representation, which was obtained using NetDRAW [31], is shown in Figure 3. The Figure shows a situation after few steps of adding edges into the structure have been done. Users are depicted as circles in the color according to group membership, apart from the reds that are members of nongroup. Every group has also a leading member that has additional responsibilities, e.g. outer group communication support. The leaders are shown as squares in the color corresponding to group membership. For the sake of lucidity, the edges weights are not shown.

¿From the figure it can be seen that the consistency is preserved and the vertexes are uniformly distributed into groups. Note that group corresponds to a hyperedge in our hypergraph model and that the implementation uses hypergraphs with hyperedge incidence 2.

5. Conclusions

The aim of the paper was to propose a new security model for mobile databases, semantic web and grids. The paper gave a brief overview on two separate levels of the security (cryptography and the trust) followed by a list of features specific to the target environments. Having summarized all requirements we described our proposal based on Virtual Organization model for the trust security level. Our model uses hypergraph theory as its mathematical basis, as the hypergraphs have abilities to reduce the space complexity of the model. The results obtained through the experimental implementation are given to verify the "evolution" phase of the proposed model demonstrated that it does not degenerate to any of the limiting cases. Although the model is based on hypergraphs with full cardinality of hyperedges, the experimental application is based on hypergraph with hyperedges' cardinalities reduced to 2.

References

[1] S. DasBit and S. Mitra, "Challenges of computing in mobile cellular environment a survey", *Elsevier B.V.*, 2003.

[2] Y. Lu, B. Bhargava, W. Wang, Y. Zhong and X. Wu, "Secure Wireless Network with Movable Base Stations", *IEICE Trans. Community*, vol. E86-B, 2003.

[3] Y. Zong, B. Bhargava and M. Mahoui, "Trustworthiness Based Authorization on WWW", *IEEE Workshop on Security in Distributed Data Warehousing*, 2001.

[4] P. K. Behera, P. K. Meher, "Prospects of Group-Based Communication in Mobile Ad hoc Networks", *Springer-Verlag Berlin Heidelberg*, 2002.

[5] A. Flaxman, A. Frieze, E. Upfal, "Efficient communication in an ad-hoc network", *Elsevier*, 2004.

[6] S. Basagni, "Remarks on Ad Hoc Networking", *Springer-Verlag*, Berlin Heidelberg, 2002.

[7] R. Molva, P. Michiardi, "Security in Ad Hoc Network", *IFIP International Federation for Information Processing*, 2003.

[8] "http://www.daml.org/"

[9] "http://www.w3.org/TR/webont-req/"

[10] J. Steiner, C . Neuman, and J.I . Schiller, "Kerberos : An Authentication Service for Open Network Systems, " *in Proc. Winter USENIX Conference*, Dallas (1988).

[11] Philip R. Zimmermann. "The Official PGP User's Guide". *MIT Press*, Boston, 1995.

[12] Carl M. Ellison, Bill Frantz, Butler Lampson, Ron Rivest, Brian M. Thomas, and Tatu Ylonen. "SPKI certificate theory," *Internet RFC 2693*. October 1999.

[13] Arsenault, A. and Turner, S. "Internet X.509 public key infrastructure: roadmap internet draft", *draft-ietf-pkix-roadmap-09.txt*, July 2002. http://www.ietf.org/internet-drafts/draft-ietf-pkix-roadmap-09.txt.

[14] P. Bonatti and P. Samarati, "Regulating service access and information release on the web.", *In CCS ´00: Proceedings of the 7th ACM conference on computer and communications security*, pages 134–143. ACM Press, 2000.

[15] N. Li and J. Mitchell, "A Role-based Trust-management Framework.", *In DARPA Information Survivability Conference and Exposition (DISCEX)*, Washington, D.C., Apr. 2003.

[16] R. Gavriloaie, W. Nejdl, D. Olmedilla, K. E. Seamons, and M. Winslett, "No registration needed: How to use declarative policies and negotiation to access sensitive resources on the semantic web.", *In 1st European Semantic Web Symposium (ESWS 2004)*, volume 3053 of Lecture Notes in Computer Science, pages 342-356, Heraklion, Crete, Greece,Springer, May 2004.

[17] M. Y. Becker and P. Sewell, "Cassandra: distributed access control policies with tunable expressiveness.", *In 5th IEEE International Workshop on Policies for Distributed Systems and Networks*, Yorktown Heights, June 2004.

[18] P. A. Bonatti and D. Olmedilla, "Driving and monitoring provisional trust negotiation with metapolicies.", *In 6th IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY 2005)*, pages 14-23, Stockholm, Sweden, Jun 2005. IEEE Computer Society.

[19] J. Basney, W. Nejdl, D. Olmedilla, V. Welch, and M. Winslett, "Negotiating trust on the grid.", *In 2nd WWW Workshop on Semantics in P2P and Grid Computing*, New York, USA, May 2004.

[20] K. Aberer and Z. Despotovic, "Managing trust in a peer-2-peer information system.", *In Proceedings of 10th International Conference on Information and Knowledge Management*, pages 310-317, 2001.

[21] E. Damiani, S. D. C. di Vimercati, S. Paraboschi, P. Samarati, and F. Violante, "A reputation-based approach for choosing reliable resources in peer-to-peer networks.", *In Proceedings of ACM Conference on Computer and Communications Security*, pages 202-216, 2002.

[22] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, " Eigenrep: Reputation management in p2p networks.", *In Proceedings of 12th International WWW Conference*, pages 640-651, 2003.

[23] C. Duma, N. Shahmehri, and G. Caronni, "Dynamic trust metrics for peer-topeer systems.", *In Proceedings of 2nd IEEE Workshop on P2P Data Management*, Security and Trust (in connection with DEXA ´05), August 2005.

[24] L. Kagal, T. Finin, and A. Joshi, "A policy based approach to security for the semantic web.", *In Proceedings of the 2nd International Semantic Web Conference*, Sanibel Island, Florida, USA, Oct. 2003.

[25] G. Tonti, J. M. Bradshaw, R. Jeffers, R. Montanari, N. Suri, and A. Uszok, "Semantic web languages for policy representation and reasoning: A comparison of KAoS, Rei and Ponder.", *In Proceedings of the 2nd International Semantic Web Conference*, Sanibel Island, Florida, USA, Oct. 2003.

[26] R. Spanek, M. Tuma, "Secure Grid-based Computing with Social-Network Based Trust Management in the Semantic Web", *submitted to NNW*, 2006.

[27] "OMG Specifications", *http://www.omg.org/technology/documents/spec_catalog.htm*.

[28] "Java RMI Specification", *http://java.sun.com/j2se/1.4.2/docs/guide/rmi/spec/rmiTOC.html*.

[29] "WSDL specification", *http://www.w3.org/2002/ws/desc/*.

[30] "SOAP specification", *http://www.w3.org/TR/soap12-part0/*.

[31] S.P. Borgatti, "NetDraw: Graph Visualization Software.", *Harvard: Analytic Technologies*, 2002, http://www.analytictech.com/netdraw.htm.

Combining Classifiers

Post-Graduate Student:
ING. DAVID ŠTEFKA

Department of mathematics
Faculty of nuclear science and physical engineering
ČVUT
Trojanova 13
120 00 Praha 2

Czech Republic

david.stefka@gmail.com

Supervisor:
ING. RNDR. MARTIN HOLEŇA, CSC.

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 2
182 07 Praha 8

Czech Republic

martin@cs.cas.cz

Field of Study:
Mathematical Engineering

Abstract

Combining classifiers has been given more and more importance in the recent years. The experiments show that a team of several different classifiers can perform better than any individual classifier in the team. This is why the combination of multiple classifiers can be used to improve the quality of classification. This arcticle tries to summarize and describe the commonly used approaches found in the literature and to give a basic overview over this field.

1. Introduction

Combination of multiple classifiers is an effective method for improving the quality (i.e. accuracy, precision, separability, sensitivity etc.) of classification. Whenever we are solving a classification task, we can construct several different classifiers, let them predict independently, and then combine their outputs somehow. The combination can be such that the strong points of the classifiers are amplified, while the deficiencies are reduced. After the combination, the final classifier can predict better than any of the single classifier in the team. This idea is known in the literature under different names, such as ”classifier combining”, ”classifier aggregation”, ”classifier fusion”, ”classifier selection”, ”mixture of experts”, etc. Basically, there are two main types of classifier combination:

- **classifier selection**, where we use some rule to determine which classifier to use for the current pattern; only this ”expert” classifier is then used for the final decision
- **classifier aggregation**, where all the classifiers in the team are used to construct the final decision

In this paper, we will deal with classifier aggregation only. A special case of classifier aggregation is combining of several classifiers of the same type, which differ only in their parameters or use different training sets. This approach is known under the name *ensemble methods* in the literature.

The paper is structured as follows: in Section 2, we formalize the term classifier and introduce some models that can be used for classifier combination, Section 3 describes some of the commonly used methods for creating an ensemble of classifiers, and in Section 4 we describe some commonly used aggregation rules. Section 5 then concludes the paper.

2. Classification

Classification is a process of dividing objects (or patterns) from some input space $\mathcal{X}$ into several disjoint sets, called classes. For example, we may try to recognize handwritten letters using a computer programme.

The software (or mathematical mapping) which realizes this task is called a classifier. We can formally describe a classifier as a mapping

$$\phi : \mathcal{X} \to \{1, \ldots, N\}, \tag{1}$$

where $\mathcal{X} \subseteq \mathcal{R}^n$ is the n-dimensional feature space and $\{1, \ldots, N\}$ is the set of class labels (we denote the respective classes by $C_1, \ldots, C_N$). An optimal, error-free classifier should also satisfy

$$\forall \mathbf{x} \in \mathcal{X} : \phi(\mathbf{x}) = i \iff x \in C_i \tag{2}$$

Construction of a classifier such that (2) holds is nearly impossible, so this condition is usually violated. Some requirements on train or test error of the classifier are used instead. To train the classifier to correctly classify previously unseen patterns, patterns from training set $\mathcal{T} \subseteq \mathcal{X}$ are used. Set $\mathcal{T}$ contains patterns, which are already classified correctly by an expert and we want our classifier to somehow generalize this knowledge to unseen patterns from $\mathcal{X}$.

Most widely-used classification methods are for example neural networks, decision trees, Bayesian classifier, k-nearest neighbour (k-NN), or support vector machines (SVM). All these methods can provide a succesfull approach to classification, but sometimes higher quality of the classification is needed. If we decide to combine different classifiers to improve the quality, we see that classifiers defined using (1) give us too little information. We could use a information of the type "what is the 'weight of classification' to each class?". To realize this, we define a classifier as a mapping

$$\phi : \mathcal{X} \to [0, 1]^N, \tag{3}$$

where $\phi(\mathbf{x}) = (\mu_1(\mathbf{x}), \ldots, \mu_N(\mathbf{x}))$ is the vector of 'weights of classification' of $\mathbf{x}$ to each class.

This type of classifier is usually denoted (for example in [1]) *measurement* classifier, while (1) is usually called *crisp* classifier. Sometimes other types of classifiers are used, for example *rank* classifier ([1]), which outputs a sorted list of all possible class labels (i.e. the most "probable" class is first, the second most "probable" is second and so on), or *possibilistic* classifier (see for example [2]), for which $\mu_i(\mathbf{x}) \in [0, \infty)$. In this paper, whenever we speak of a classifier, we mean measurement classifier.

Measurement classifier can be converted to crisp and vice versa:

$$\phi_{crisp}(\mathbf{x}) = \arg\max\{\mu_i(\mathbf{x}) | i = 1, \ldots, N\} \tag{4}$$

$$\phi_{measurement}(\mathbf{x}) = (\delta(1, \phi(\mathbf{x})), \ldots, \delta(N, \phi(\mathbf{x}))),$$

where $\delta(i, \phi(\mathbf{x})) = 1$ if $\phi(\mathbf{x}) = i$, 0 otherwise.

The interpretation of $\phi(\mathbf{x})$ depends mainly on the classifier used; possible interpretations can be:

- **posterior probability**, $\mu_i = P(\mathbf{x} \in C_i | \mathbf{x})$, for example for Bayesian classifier
- $\mu_i =$ degree of **membership** of $\mathbf{x}$ to a fuzzy set C_i
- $\mu_i =$ normalized **distance** to the nearest representant of class i, for 1-NN classifier
- $\mu_i =$ **number of nearest representants** of class i (divided by k), for k-NN classifier
- ...

Sometimes additional conditions on μ_i are demanded, for example $\sum_i \mu_i(\mathbf{x}) = 1$ for probabilistic interpretation of $\phi(\mathbf{x})$. However, these conditions are not important for classifier combining itself, so we only require $\mu_i \in [0, 1]$. Since we aim to present a general approach to classifier combining, we also do not care about the interpretation of μ_i, and we only interpret these values as some general 'weights of classification' to class C_i.

Bagging Create an ensemble of classifiers $\phi_1, \ldots, \phi_k$ using training set $\mathcal{T} = \{\mathbf{x}_i \in \mathcal{X} | i = 1, \ldots, m\}$:

1. Set $j = 1$.
2. Sample m-times with replacement from the training set $\mathcal{T}$, using uniform distribution, i.e. $P(\mathbf{x}_i) = 1/m \; \forall i$, creating a new training set $\mathcal{T}_j$.
3. Create a classifier ϕ_j using training set $\mathcal{T}_j$.
4. If $j < k$, increment j and return to (2), otherwise end with output $\phi_1, \ldots, \phi_k$.

Figure 1: The bagging algorithm

3. Ensemble methods

It is straightforward that if we want to build a team of several classifiers, these classifiers should all perform well. However, we should care not only about the train and test error of the classifiers in the team, but also about the diversity of the classifiers. A team of mutually independent and diverse classifiers can perform quite well in comparison to any of the single classifiers. However, if the classifiers in the team are "negatively dependent", i.e. they make dependent errors (when one of the classifiers predicts wrong, then also many of the other classifiers predict wrong), the team of the classifiers can be worse than the single best classifier. On the other hand, the team can be such that only a few classifiers in the team predict wrong and the rest of the team predicts well for each pattern. In that case, classifier combining can significantly improve quality of the prediction over the single best classifier.

Of course, diversity of the team is not the only quality determining the improvement of the performance, but it is an important feature. The diversity of a classifier team can be measured by various measures. Many diversity measures are analyzed and compared to each other in [3]. Also the role of diversity of the team on the final error rate is discussed here. An entropy based diversity measure is introduced in [4]. The role of diversity on the quality of the final aggregated classifier is still a subject of research.

Several methods for building a diverse team of classifiers were developed. Among them, ensemble methods are used frequently. These methods construct a set of classifiers of the same type, which differ only in their parameters or have different training sets. Here we will describe bagging and boosting, which work with modifying the training set. Other methods, constructing for example k-NN classifiers with different metrics ([5]), SVM with different kernels, or neural networks with different architectures, can also be developed.

The idea of bagging (Bootstrap AGGregatING) was introduced by Breiman in [6]. Bagging takes bootstrap samples (random uniformly distributed sample with replacement of the same size as the training set) and uses this set as a new training set for the individual classifier. Some patterns are selected more than once, some of them are not present in the new training set. The bagging algorithm is summarized in Fig. 1.

Another approach, presented for example in [7], is called boosting. Boosting is a general machine learning method for improving the performance of a weak (i.e. simple, not complex) classifier, which works by repeatedly running a given weak algorithm with different distributions over the training set. In this paper, we present the AdaBoost algorithm by Freund and Schapire ([7]). AdaBoost also uses bootstrap technique, with the only difference that the probability of selecting a pattern is not uniform and also changes in time. The first classifier is created by uniform bootstrap sampling, then its train error rate (using the original training set) is measured and those patterns which are classified incorrectly are given greater probability of selection. If the train error of the classifier is too high (for example higher than 0.5 for binary classifier), the classifier is discarded and the probabilities are set back to uniform. Then the whole process is repeated. After all classifiers are constructed, the aggregation is usually done using a weighted average, where the individual weights correspond to the appropriate train errors. Of course, any other aggregation rule can be used.

AdaBoost.M1 Create an ensemble of classifiers $\phi_1, \ldots, \phi_k$ using training set $\mathcal{T} = \{\mathbf{x}_i \in \mathcal{X} | i = 1, \ldots, m\}$; for each training pattern $\mathbf{x}_i$ we know the corresponding class label y_i:

1. Set $j = 1$.
2. Initialize the discrete probabilities $P_j(\mathbf{x}_i) = 1/m \; \forall i$.
3. Sample m-times with replacement from the training set $\mathcal{T}$, using the discrete distribution P_j, creating a new training set $\mathcal{T}_j$.
4. Create a classifier ϕ_j with training set $\mathcal{T}_j$.
5. Calculate the weighted error of ϕ_j:
$$err_j = \sum_{i:\phi(\mathbf{x}_i) \neq y_i} P_j(\mathbf{x}_i)$$
6. If $err_j \geq 0.5$, discard the classifier ϕ_j and return to (2).
7. Set $\beta_j = err_j/(1 - err_j)$.
8. Update P:
$$P_{j+1}(\mathbf{x}_i) = \frac{P_j(\mathbf{x}_i)}{C} \cdot \begin{cases} \beta_j & for\ \phi(\mathbf{x}_i) = y_i \\ 1 & otherwise \end{cases},$$
where $C = \sum_i P_j(\mathbf{x_i})$ is a normalizing constant.
9. If $j < k$, increment j and return to (3), otherwise end with output $\phi_1, \ldots, \phi_k$. If weighted average is used for aggregation of the classifiers, use the weights $\omega_j = \log \frac{1}{\beta_j}$.

Figure 2: The AdaBoost.M1 algorithm

Here we present the AdaBoost.M1 algorithm, which can be used for binary classification. In [7], there is also described a modification for multi-class classification, AdaBoost.M2. These two methods are the same for binary classification and differ only in the handling of problems with three or more classes; for the sake of simplicity, we will describe AdaBoost.M1 only. It is described in Fig. 2.

Boosting and bagging are experimentally compared in [7], where in the most cases boosting performs better than bagging. In [8], theoretical aspects of bagging and boosting are dealt with and a new class of methods, called arcing (the special case of which is AdaBoost), is introduced. Also a decomposition of classifier error to bias and variance and the effect of reducing variance by classifier combining is discussed here.

4. Classifier aggregation

After we have constructed an ensemble consisting of k classifiers $\phi_1, \ldots, \phi_k$, we need to aggregate them to get the final classifier. The output of the ensemble can be structurred to a $k \times N$ matrix:

$$\begin{pmatrix} \phi_1(\mathbf{x}) \\ \phi_2(\mathbf{x}) \\ \vdots \\ \phi_k(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \mu_{1,1}(\mathbf{x}) & \mu_{1,2}(\mathbf{x}) & \ldots & \mu_{1,N}(\mathbf{x}) \\ \mu_{2,1}(\mathbf{x}) & \mu_{2,2}(\mathbf{x}) & \ldots & \mu_{2,N}(\mathbf{x}) \\ & & \ddots & \\ \mu_{k,1}(\mathbf{x}) & \mu_{k,2}(\mathbf{x}) & \ldots & \mu_{k,N}(\mathbf{x}) \end{pmatrix}, \tag{5}$$

where $\mu_{i,j}(\mathbf{x})$ is the 'weight of classification' of $\mathbf{x} \in C_j$ assigned by classifier ϕ_i. In [3], this matrix is referred to as a *decision profile*.

To aggregate the results of the classifiers, we use some aggregation rule $\mathcal{F}$:

$$\Phi(\mathbf{x}) = \mathcal{F}(\phi_1(\mathbf{x}), \ldots, \phi_k(\mathbf{x})), \tag{6}$$

where Φ is the final classifier. Usually, Φ is a measurement classifier, and for the final prediction, its output is converted to crisp answer using (4).

A very good overview and experimental comparison of different aggregation rules can be found in [2]. We will present some of the commonly used aggregation rules found in the literature. The rules can be divided into four main types — arithmetic, probabilistic, fuzzy, and hierarchical classification. Also other methods, for example Dempster-Shafer fusion ([2], [9]) can be used.

4.1. Arithmetic rules

The simplest rules use only some simple arithmetic operations to combine the classifier outputs. We can use for example:

- **voting** (for crisp classifiers) – each classifier has one vote, and the class with maximal votes is taken for result. Ties are broken arbitrarily.
- **maximum** – for each column of the matrix (5), we use the maximal value only, in other words:

$$\mu_j(\mathbf{x}) = \max\{\mu_{i,j}(\mathbf{x}) | i = 1, \ldots, k\},\ j = 1, \ldots, N$$

- **minimum** – same as the previous, only the minimal value is chosen instead:

$$\mu_j(\mathbf{x}) = \min\{\mu_{i,j}(\mathbf{x}) | i = 1, \ldots, k\},\ j = 1, \ldots, N$$

- **mean** – this approach tries to approximate the "typical weight of classification" for each class through all the classifiers:

$$\mu_j(\mathbf{x}) = \frac{1}{k} \sum_{i=1}^{k} \mu_{i,j}(\mathbf{x}),\ j = 1, \ldots, N$$

- **weighted mean** – suppose that we have defined a weight ω_i for each classifier (for example the weights from AdaBoost.M1 algorithm). Using these weights, we can add more importance to the classifiers that predict better than the others:

$$\mu_j(\mathbf{x}) = \frac{\sum_{i=1}^{k} \omega_i \mu_{i,j}(\mathbf{x})}{\sum_{i=1}^{k} \omega_i},\ j = 1, \ldots, N$$

- **product** – "weight of classification" to class j is computed as a product of the ith column of matrix (5):

$$\mu_j(\mathbf{x}) = \prod_{i=1}^{k} \mu_{i,j}(\mathbf{x}),\ j = 1, \ldots, N \tag{7}$$

4.2. Probabilistic rules

Here we present two different aggregation rules which use probability theory to combine the classifiers – behavior knowledge space, and product rule. The former creates all possible combinations of crisp class labels, and in each combination, the most probable class is selected. The latter tries to estimate the posterior probability that the current pattern belongs to jth class for a given j.

- **behavior knowledge space (BKS)** – this approach works with crisp classifiers. Every possible combination of crisp outputs is generated, and in each partition (combination of the crisp outputs), the most frequent class (among patterns from the training set, for which the ensemble classifiers give the

same combination of crisp outputs) is taken for the final classification of the team. If in any partition there are no training patterns, or tie occurs, a random choice (using uniform distribution) among the appropriate classes is made. For example, if we have $k = 5$ classifiers and $N = 3$ classes, the possible crisp outputs are:

$$[1,1,1,1,1][1,1,1,1,2][1,1,1,1,3][1,1,1,2,1]\dots[3,3,3,3,3]$$

If in the first partition $[1,1,1,1,1]$(which means that every classifier predicts class C_1) of the training set (i.e. patterns, for which every classifier predicts class C_1), there are r, s, t representants of the classes C_1, C_2 and C_3 respectively, $\arg\max_{C_1,C_2,C_3}\{r,s,t\}$ is the final result in this partition. Ties are broken at random (using uniform distribution).

- **product** – if we can interpret the outputs of the ensemble classifiers as estimates of posterior probabilities, i.e. $\mu_{i,j}(\mathbf{x}) = P_i(\mathbf{x} \in C_j|\mathbf{x})$, and the classifiers $\phi_1, \dots, \phi_k$ are mutually independent, then the product aggregation rule (7) has also probabilistic interpretation (conjunction of all the probabilities). We can also take into account the prior probabilities of the classes $P(C_1), \dots, P(C_N)$, $P(C_j) = \frac{\#\{\mathbf{x}\in\mathcal{T}|\mathbf{x}\in C_j\}}{\#\mathcal{T}}$, resulting in the final Bayes approximation:

$$\mu_j(\mathbf{x}) = \frac{\prod_{i=1}^{k}\mu_{i,j}(\mathbf{x})}{P(C_j)^{k-1}},\ j = 1,\dots,N$$

4.3. Fuzzy rules

If we can interpret the classes $C_1, \dots, C_N$ as fuzzy sets and the classifier outputs $\mu_i(\mathbf{x})$ as fuzzy membership of $\mathbf{x}$ to class C_i, we can use fuzzy logic approaches to aggregate the outputs. Comparison of fuzzy versus nonfuzzy methods for classifier combining can be found in [10].

- **fuzzy integral** – let the universal set be the set of all classifiers $\mathcal{U} = \{\phi_1, \dots, \phi_k\}$, g fuzzy measure on $\mathcal{U}$ (representing the importance, or quality, of the classifiers), and $A_j(\mathbf{x}) = (\mu_{1,j}(\mathbf{x}), \dots, \mu_{k,j}(\mathbf{x}))^T$ the jth column of matrix (5) (representing the membership of $\mathbf{x}$ to class j for all the classifiers). $A_j(\mathbf{x})$ can be treated as a fuzzy set on $\mathcal{U}$. The fuzzy measure g can be computed from k point-wise values of the measure $g^{(1)}, \dots, g^{(k)}$ (called fuzzy densities). We can then use Sugeno or Choquet fuzzy integral to aggregate $A_j(\mathbf{x})$ using the fuzzy measure g. The aggregation using fuzzy integral (as described in [10]) is ilustrated in Fig. 3.

- **decision templates** – this technique was developed by Kuncheva et al. in [2]. For each class C_j, a decision template (which expresses a "typical" ensemble output for this class) is defined as a $k \times N$ matrix DT, which elements are averages of the ensemble outputs through all the training patterns belonging to C_j:

$$(DT_j)_{r,s} = \frac{\sum_{\mathbf{x}\in\mathcal{T}:\mathbf{x}\in C_j}\mu_{r,s}(\mathbf{x})}{\#\{\mathbf{x}\in\mathcal{T} : \mathbf{x}\in C_j\}} \tag{8}$$

  We interpret the matrices (5) and DT_j as fuzzy sets on a universal set $\mathcal{U}$ of $k \times N$ elements. When $\mathbf{x}$ is submitted for classification, ensemble output (5) is computed and some kind of similarity of this output and decision template for each class are measured. These similarities can be for instance fuzzy similarities of these sets, fuzzy inclusion indices, consistency measures or any matrix norm of the difference of the two matrices. In [2], 11 different similarities are proposed, and also the effect of using different "similarity" measures is discussed. According to the experimental results in [2] and [10], we can recommend the Euclidean distance, or the S_3 similarity measure: let A, B be fuzzy sets on $\mathcal{U}$, and μ_A, μ_B their membership functions. Then we define

$$S_3(A,B) = 1 - ||A \triangle B||,$$

  where $||A||$ denotes the relative cardinality of fuzzy set A and $\mu_{A\triangle B}(u) = \max\{\mu_{A\cap\bar{B}}(u), \mu_{\bar{A}\cap B}(u)\}$, where $\cap$ is a fuzzy intersection and $\bar{\cdot}$ is a fuzzy complement.

Fuzzy integral

1. Fix the k fuzzy densities $g^{(1)}, \ldots, g^{(k)}$. These values can be for example estimates of the classifier accuracies.
2. Calculate λ as the only real root greater than -1 of the equation:
$$\lambda + 1 = \prod_{i=1}^{k} (1 + \lambda g^{(i)})$$
3. When $\mathbf{x}$ is submitted for classification, perform steps (4)-(7) for $j = 1, \ldots, N$.
4. Sort the jth column of matrix (5) in ascending order, obtaining the membership values $\mu_{i_1,j}(\mathbf{x}), \ldots, \mu_{i_k,j}(\mathbf{x})$.
5. Sort the fuzzy densities correspondingly: $g^{(i_1)}, \ldots, g^{(i_k)}$ and set $g(1) = g^{(i_1)}$.
6. For each $t = 2, \ldots, k$ calculate
$$g(t) = g^{(i_t)} + g(t-1) + \lambda g^{(i_t)} g(t-1)$$
7. (a) For Sugeno integral, the final value is
$$\mu_j(\mathbf{x}) = \max_{t=1,\ldots,k} \{\min\{\mu_{i_t,j}(\mathbf{x}), g(t)\}\}$$
   (b) For Choquet integral, the final value is
$$\mu_j(\mathbf{x}) = \mu_{i_1,j}(\mathbf{x}) + \sum_{t=2}^{k} \left(\mu_{i_{t-1},j}(\mathbf{x}) - \mu_{i_t,j}(\mathbf{x})\right) g(t-1)$$

Figure 3: Fuzzy integral

4.4. Hierarchical classification

The output (5) of the ensemble can be treated as a vector of features $\in [0, 1]$ of length kN. We can use this space as an input for a top-level classifier and train a crisp classifier $\Phi : [0, 1]^{kN} \to \{1, \ldots, N\}$. This approach has one drawback, that if the classifiers $\phi_1, \ldots, \phi_k$ predict "nearly crisp labels" (i.e. only one member of the vector $\phi(\mathbf{x})$ is close to one and the remaining are close to zero), then the covariance matrices of the training patterns are close to singular. Therefore, whenever we use normal distribution for modelling the pattern distribution in the classes (for example in Bayesian classifier, linear and quadratic discriminant analysis and many others), the results may be inaccurate.

5. Conclusion

In this paper, we described different approaches to classifier combination. Possible mathematical models of classifiers that can be used for classifier combining were introduced. Also two methods for creating an ensemble of classifiers, bagging and boosting, were described. Then some of the commonly used methods for aggregation of the outputs of multiple classifiers were described. In the future, I would like to study the fuzzy integral aggregation rule, and also try to use other types of fuzzy integral than Sugeno and Choquet type. I would also like to try aggregating the classifiers using a Mamdani-Assilian fuzzy controller.

References

[1] O. Melnik, Y. Vardi and C.-H. Zhang, "Mixed group ranks: Preference and confidence in classifier combination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 8, pp. 973–981, 2004.

[2] L. I. Kuncheva, J. C. Bezdek and R. P. W. Duin, "Decision templates for multiple classifier fusion: an experimental comparison.," *Pattern Recognition*, vol. 34, no. 2, pp. 299–314, 2001.

[3] L. I. Kuncheva and C. J. Whitaker, "Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy," *Machine Learning*, vol. 51, no. 2, pp. 181–207, 2003.

[4] W. Liu, Z. Wu and G. Pan, "An entropy-based diversity measure for classifier combining and its application to face classifier ensemble thinning," in *Sinbiometrics*, pp. 118–124, 2004.

[5] Y. Bao, N. Ishii and X. Du, "Combining multiple k-nearest neighbor classifiers using different distance functions.," in *IDEAL*, pp. 634–641, 2004.

[6] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[7] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *International Conference on Machine Learning*, pp. 148–156, 1996.

[8] L. Breiman, "Bias, variance, and arcing classifiers," tech. rep., Technical Report 460, Statistics Department, University of California, 1996.

[9] M. R. Ahmadzadeh and M. Petrou, "Use of dempster-shafer theory to combine classifiers which use different class boundaries," *Pattern Analysis & Applications*, vol. 6, pp. 41–46, 2003.

[10] L. I. Kuncheva, "Fuzzy versus nonfuzzy in combining classifiers designed by boosting," *IEEE Transactions on Fuzzy Systems*, vol. 11, no. 6, pp. 729–741, 2003.

Ústav Informatiky AV ČR
DOKTORANDSKÝ DEN '06

Vydal
MATFYZPRESS
vydavatelství
Matematicko-fyzikální fakulty
Univerzity Karlovy
Sokolovská 83, 186 75 Praha 8
jako svou 181. publikaci

Obálku navrhl František Hakl

Z předloh připravených v systému LaTeX
vytisklo Reprostředisko MFF UK
Sokolovská 83, 186 75 Praha 8

Vydání první

Praha 2006

ISBN 80-86732-87-8